

OC Blaberidae; Leucophaea.  
 OX NCBI\_TaxID=6988;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Head;  
 RA Holman G.M., Cook B.J., Nachman R.J.;  
 RT "Isolation, primary structure and synthesis of two neuropeptides  
 RT from Leucophaea maderae: members of a new family of  
 RT Cephalomyotropins.";  
 RL Comp. Biochem. Physiol. 84C:205-211(1986).  
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE  
 CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).  
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 8 8 AMIDATION.  
 SQ SEQUENCE 8 AA; 893 MW; DC6365B449CDC76A CRC64;

Query Match 25.0%; Score 2; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WG 5  
 ||  
 Db 7 WG 8

#### RESULT 10

##### LCK2\_LEUMA

ID LCK2\_LEUMA STANDARD; PRT; 8 AA.  
 AC P21141;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 01-MAY-1991 (Rel. 18, Last annotation update)  
 DE Leucokinin II (L-II).  
 OS Leucophaea maderae (Madeira cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;  
 OC Blaberidae; Leucophaea.  
 OX NCBI\_TaxID=6988;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Head;  
 RA Holman G.M., Cook B.J., Nachman R.J.;  
 RT "Isolation, primary structure and synthesis of two neuropeptides  
 RT from Leucophaea maderae: members of a new family of  
 RT Cephalomyotropins.";  
 RL Comp. Biochem. Physiol. 84C:205-211(1986).  
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE  
 CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).  
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 8 8 AMIDATION.  
 SQ SEQUENCE 8 AA; 852 MW; DC6365A5B9C8676A CRC64;

Query Match 25.0%; Score 2; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy            4 WG 5  
               ||  
 Db            7 WG 8

RESULT 11

LCK3\_LEUMA

ID LCK3\_LEUMA            STANDARD;            PRT;            8 AA.  
 AC P21142;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 01-MAY-1991 (Rel. 18, Last annotation update)  
 DE Leucokinin III (L-III).  
 OS Leucophaea maderae (Madeira cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;  
 OC Blaberidae; Leucophaea.  
 OX NCBI\_TaxID=6988;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Head;  
 RA Holman G.M., Cook B.J., Nachman R.J.;  
 RT "Primary structure and synthesis of two additional neuropeptides  
 RT from Leucophaea maderae: members of a new family of  
 RT Cephalomyotropins.";  
 RL Comp. Biochem. Physiol. 84C:271-276(1986).  
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE  
 CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).  
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES            8            8            AMIDATION.  
 SQ SEQUENCE            8 AA;    910 MW;    DC6365B449C866DA CRC64;

Query Match            25.0%;    Score 2;    DB 1;    Length 8;  
 Best Local Similarity    100.0%;    Pred. No. 1.3e+05;  
 Matches            2;    Conservative            0;    Mismatches            0;    Indels            0;    Gaps            0;

Qy            4 WG 5  
               ||  
 Db            7 WG 8

RESULT 12

LCK4\_LEUMA

ID LCK4\_LEUMA            STANDARD;            PRT;            8 AA.  
 AC P21143;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 01-MAY-1991 (Rel. 18, Last annotation update)  
 DE Leucokinin IV (L-IV).  
 OS Leucophaea maderae (Madeira cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;  
 OC Blaberidae; Leucophaea.  
 OX NCBI\_TaxID=6988;  
 RN [1]

RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Head;  
 RA Holman G.M., Cook B.J., Nachman R.J.;  
 RT "Primary structure and synthesis of two additional neuropeptides  
 from *Leucophaea maderae*: members of a new family of  
 RT Cephalomyotropins.";  
 RL Comp. Biochem. Physiol. 84C:271-276(1986).  
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE  
 CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).  
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 8 8 AMIDATION.  
 SQ SEQUENCE 8 AA; 906 MW; DC6365B1E9D5BDDA CRC64;

Query Match 25.0%; Score 2; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WG 5  
 ||  
 Db 7 WG 8

#### RESULT 13

##### LCK5\_LEUMA

ID LCK5\_LEUMA STANDARD; PRT; 8 AA.  
 AC P19987;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Leucokinin V (L-V).  
 OS *Leucophaea maderae* (Madeira cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;  
 OC Blaberidae; *Leucophaea*.  
 OX NCBI\_TaxID=6988;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Head;  
 RX MEDLINE=87052651; PubMed=2877794;  
 RA Holman G.M., Cook B.J., Nachman R.J.;  
 RT "Isolation, primary structure, and synthesis of leucokinins V and VI:  
 RT myotropic peptides of *Leucophaea maderae*.";  
 RL Comp. Biochem. Physiol. 88C:27-30(1987).  
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE  
 CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).  
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.  
 DR PIR; JS0315; JS0315.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 8 8 AMIDATION.  
 SQ SEQUENCE 8 AA; 784 MW; 736365A5B9C865B8 CRC64;

Query Match 25.0%; Score 2; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WG 5

Db

||  
7 WG 8

RESULT 14

LCK6\_LEUMA

ID LCK6\_LEUMA STANDARD; PRT; 8 AA.  
AC P19988;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Leucokinin VI (L-VI).  
OS Leucophaea maderae (Madeira cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;  
OC Blaberidae; Leucophaea.  
OX NCBI\_TaxID=6988;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Head;  
RX MEDLINE=87052651; PubMed=28777794;  
RA Holman G.M., Cook B.J., Nachman R.J.;  
RT "Isolation, primary structure, and synthesis of leucokinins V and VI:  
RT myotropic peptides of Leucophaea maderae.";  
RL Comp. Biochem. Physiol. 88C:27-30(1987).  
CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE  
CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).  
CC -!- SIMILARITY: TO THE OTHER LEUCOKININS, AND TO MANDUCA SEXTA AND  
CC HELIOTHIS ZEA ADIPOKINETIC HORMONE.  
DR PIR; JS0316; JS0316.  
KW Neuropeptide; Amidation; Pyrrolidone carboxylic acid.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 8 8 AMIDATION.  
SQ SEQUENCE 8 AA; 935 MW; 9D6365B1E9D5A5A6 CRC64;

Query Match 25.0%; Score 2; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WG 5

||

Db 7 WG 8

RESULT 15

LCK7\_LEUMA

ID LCK7\_LEUMA STANDARD; PRT; 8 AA.  
AC P19989;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 01-FEB-1991 (Rel. 17, Last annotation update)  
DE Leucokinin VII (L-VII).  
OS Leucophaea maderae (Madeira cockroach).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;  
OC Blaberidae; Leucophaea.  
OX NCBI\_TaxID=6988;

RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Head;  
 RA Holman G.M., Cook B.J., Nachman R.J.;  
 RT "Isolation, primary structure and synthesis of leucokinins VII and  
 RT VIII: the final members of this new family of cephalomyotropic  
 RT peptides isolated from head extracts of *Leucophaea maderae*";  
 RL Comp. Biochem. Physiol. 88C:31-34(1987).  
 CC -!- FUNCTION: THIS CEPHALOMYOTROPIC PEPTIDE STIMULATES CONTRACTILE  
 CC ACTIVITY OF COCKROACH PROTODEUM (HINDGUT).  
 CC -!- SIMILARITY: TO THE OTHER LEUCOKININS.  
 DR PIR; JS0317; JS0317.  
 KW Neuropeptide; Amidation.  
 FT MOD\_RES 8 8 AMIDATION.  
 SQ SEQUENCE 8 AA; 866 MW; DC6365A5B9CDC76A CRC64;

Query Match 25.0%; Score 2; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 WG 5  
 ||  
 Db 7 WG 8

Search completed: November 13, 2003, 10:34:00  
 Job time : 5.33333 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:58:36 ; Search time 19.6667 Seconds  
 (without alignments)  
 104.971 Million cell updates/sec

Title: US-09-228-866-8  
 Perfect score: 8  
 Sequence: 1 CLDWGRIC 8

Scoring table: OLIGO  
 Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 7516

Minimum DB seq length: 7  
 Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : SPTREMBL\_23:\*

```

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	4	50.0	17	2	006946	006946	salmonella
2	3	37.5	8	2	032560	032560	escherichia
3	3	37.5	9	8	Q8MBF4	Q8mbf4	ipomoea qua
4	3	37.5	10	8	Q8MAZ9	Q8maz9	dicranostyl
5	3	37.5	10	8	Q8MBB7	Q8mbb7	merremia ae
6	3	37.5	11	8	Q8MAZ1	Q8maz1	maripa pani
7	3	37.5	11	8	Q8MB39	Q8mb39	wilsonia hu
8	3	37.5	11	8	Q8MB58	Q8mb58	seddera hir
9	3	37.5	11	8	Q8MAZ3	Q8maz3	maripa repe
10	3	37.5	11	8	Q8MBE1	Q8mbe1	ipomoea alb
11	3	37.5	11	8	Q8MB77	Q8mb77	odonellia h
12	3	37.5	11	8	Q8MB79	Q8mb79	aniseia arg
13	3	37.5	11	8	Q8MB97	Q8mb97	merremia pe
14	3	37.5	12	7	Q31006	Q31006	bos taurus
15	3	37.5	12	8	Q8MAX7	Q8max7	tridynamia
16	3	37.5	14	2	Q93CI1	Q93ci1	escherichia
17	3	37.5	14	4	Q9UHM5	Q9uhm5	homo sapien
18	3	37.5	14	10	P82327	P82327	pisum sativ
19	3	37.5	15	2	Q47892	Q47892	freymella d
20	3	37.5	15	4	Q9BQT3	Q9bqt3	homo sapien
21	3	37.5	15	11	Q9QUW3	Q9quw3	rattus sp.
22	3	37.5	15	12	Q86576	Q86576	subterranea
23	3	37.5	16	4	Q9UCJ7	Q9ucj7	homo sapien
24	3	37.5	16	8	Q8HTT0	Q8htt0	columnea sp
25	3	37.5	16	8	Q8HTS9	Q8hts9	pinguicula
26	3	37.5	16	8	Q8HTS8	Q8hts8	pinguicula
27	3	37.5	16	8	Q8HTS7	Q8hts7	pinguicula
28	3	37.5	16	8	Q8HTS6	Q8hts6	utricularia
29	3	37.5	16	8	Q8HTS5	Q8hts5	utricularia
30	3	37.5	16	8	Q8HTS4	Q8hts4	utricularia

31	3	37.5	16	9	Q8H9Z9	Q8h9z9 bacterioph
32	3	37.5	16	15	Q75710	Q75710 human immun
33	3	37.5	17	2	Q9ZG32	Q9zg32 chlamydia t
34	3	37.5	17	2	P82586	P82586 streptococc
35	3	37.5	17	8	Q8HRZ4	Q8hrz4 ephedra sin
36	3	37.5	18	2	Q52411	Q52411 thermophili
37	3	37.5	18	4	Q16028	Q16028 homo sapien
38	3	37.5	18	8	Q9GE28	Q9ge28 amborella t
39	3	37.5	18	8	Q9ZY82	Q9zy82 encarsia fo
40	3	37.5	18	8	Q8HS04	Q8hs04 arabidopsis
41	3	37.5	18	8	Q8HQK6	Q8hqq6 conger myri
42	3	37.5	18	8	Q8HB92	Q8hb92 conger myri
43	3	37.5	19	2	Q9R517	Q9r517 mycobacteri
44	3	37.5	19	8	Q8HQK7	Q8hqq7 conger myri
45	3	37.5	19	8	Q8HC26	Q8hc26 conger myri

# ALIGNMENTS

## RESULT 1

O06946

ID O06946 PRELIMINARY; PRT; 17 AA.  
AC O06946;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Orf238 (Fragment).  
OS Salmonella typhimurium.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
OX NCBI\_TaxID=602;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LT2;  
RX MEDLINE=98036037; PubMed=9370270;  
RA Hayes F., Lubetzki S.A., Sherratt D.J.;  
RT "Salmonella typhimurium specifies a circular chromosome dimer  
RT resolution system which is homologous to the Xer site-specific  
RT recombination system of Escherichia coli.";  
RL Gene 198:105-110(1997).  
DR EMBL; U92525; AAC45779.1; -.  
FT NON\_TER 17 17  
SQ SEQUENCE 17 AA; 2072 MW; BEB65CE8F0F9F529 CRC64;

Query Match 50.0%; Score 4; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GRIC 8  
|||  
Db 8 GRIC 11

## RESULT 2

O32560

ID O32560 PRELIMINARY; PRT; 8 AA.

AC O32560;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Propionate kinase (Fragment).  
 GN TDCD.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=W3110;  
 RX MEDLINE=99449059; PubMed=10520749;  
 RA Hesslinger C., Sawers G.;  
 RT "The tdcE gene in Escherichia coli W3110 is separated from the rest of  
 RT the tdc operon by insertion of IS5 elements.";  
 RL DNA Seq. 9:183-188(1998).  
 DR EMBL; AJ001620; CAA04875.1; -.  
 KW Kinase.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 8 AA; 1000 MW; 3A505EB044140DC4 CRC64;

Query Match 37.5%; Score 3; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 RIC 8  
 |||  
 Db 4 RIC 6

#### RESULT 3

##### Q8MBF4

ID Q8MBF4 PRELIMINARY; PRT; 9 AA.  
 AC Q8MBF4;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE PsbJ (Fragment).  
 GN PSBJ.  
 OS Ipomoea quamoclit.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamiids; Solanales; Convolvulaceae; Ipomoea.  
 OX NCBI\_TaxID=89660;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stefanovic S., Krueger L., Olmstead R.G.;  
 RT "Monophyly of the Convolvulaceae and circumscription of their major  
 RT lineages based on DNA sequences of multiple chloroplast loci.";  
 RL Am. J. Bot. 0:0-0(2002).  
 DR EMBL; AY100854; AAM55540.1; -.  
 KW Chloroplast.  
 FT NON\_TER 9 9  
 SQ SEQUENCE 9 AA; 961 MW; DD59440861B1AAAD CRC64;



Query Match 37.5%; Score 3; DB 8; Length 9;  
Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GRI 7  
|||  
Db 6 GRI 8

RESULT 4

Q8MAZ9

ID Q8MAZ9 PRELIMINARY; PRT; 10 AA.  
AC Q8MAZ9;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE PsbJ (Fragment).  
GN PSBJ.  
OS Dicranostyles ampla.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; lamiids; Solanales; Convolvulaceae; Dicranostyles.  
OX NCBI\_TaxID=197378;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Stefanovic S., Krueger L., Olmstead R.G.;  
RT "Monophyly of the Convolvulaceae and circumscription of their major  
RT lineages based on DNA sequences of multiple chloroplast loci."  
RL Am. J. Bot. 0:0-0(2002).  
DR EMBL; AY100933; AAM55853.1; -.  
KW Chloroplast.  
FT NON\_TER 10 10  
SQ SEQUENCE 10 AA; 1074 MW; 836D59440861B1AA CRC64;

Query Match 37.5%; Score 3; DB 8; Length 10;  
Best Local Similarity 100.0%; Pred. No. 2.5e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GRI 7 .  
|||  
Db 6 GRI 8

RESULT 5

Q8MBB7

ID Q8MBB7 PRELIMINARY; PRT; 10 AA.  
AC Q8MBB7;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE PsbJ (Fragment).  
GN PSBJ.  
OS Merremia aegyptia.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamiids; Solanales; Convolvulaceae; Merremia.  
 OX NCBI\_TaxID=197413;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stefanovic S., Krueger L., Olmstead R.G.;  
 RT "Monophyly of the Convolvulaceae and circumscription of their major  
 RT lineages based on DNA sequences of multiple chloroplast loci.";  
 RL Am. J. Bot. 0:0-0(2002).  
 DR EMBL; AY100875; AAM55624.1; -.  
 KW Chloroplast.  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1074 MW; 836D59440861B1AA CRC64;

Query Match 37.5%; Score 3; DB 8; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GRI 7  
 |||  
 Db 6 GRI 8

#### RESULT 6

Q8MAZ1

ID Q8MAZ1 PRELIMINARY; PRT; 11 AA.  
 AC Q8MAZ1;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE PsbJ (Fragment).  
 GN PSBJ.  
 OS Maripa paniculata.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamiids; Solanales; Convolvulaceae; Maripa.  
 OX NCBI\_TaxID=197411;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stefanovic S., Krueger L., Olmstead R.G.;  
 RT "Monophyly of the Convolvulaceae and circumscription of their major  
 RT lineages based on DNA sequences of multiple chloroplast loci.";  
 RL Am. J. Bot. 0:0-0(2002).  
 DR EMBL; AY100937; AAM55869.1; -.  
 KW Chloroplast.  
 FT NON\_TER 11 11  
 SQ SEQUENCE 11 AA; 1260 MW; 93736D59440861B1 CRC64;

Query Match 37.5%; Score 3; DB 8; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GRI 7  
 |||  
 Db 6 GRI 8

# RESULT 7

Q8MB39

ID Q8MB39 PRELIMINARY; PRT; 11 AA.  
AC Q8MB39;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE PsbJ (Fragment).  
GN PSBJ.  
OS *Wilsonia humilis*.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; lamiids; Solanales; Convolvulaceae; *Wilsonia*.  
OX NCBI\_TaxID=197481;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Stefanovic S., Krueger L., Olmstead R.G.;  
RT "Monophyly of the Convolvulaceae and circumscription of their major  
RT lineages based on DNA sequences of multiple chloroplast loci.";  
RL Am. J. Bot. 0:0-0(2002).  
DR EMBL; AY100914; AAM55777.1; -.  
KW Chloroplast.  
FT NON\_TER 11 11  
SQ SEQUENCE 11 AA; 1260 MW; 93736D59440861B1 CRC64;

Query Match 37.5%; Score 3; DB 8; Length 11;  
Best Local Similarity 100.0%; Pred. No. 2.7e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GRI 7  
|||  
Db 6 GRI 8

# RESULT 8

Q8MB58

ID Q8MB58 PRELIMINARY; PRT; 11 AA.  
AC Q8MB58;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE PsbJ (Fragment).  
GN PSBJ.  
OS *Seddera hirsuta*.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; lamiids; Solanales; Convolvulaceae; *Seddera*.  
OX NCBI\_TaxID=197444;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Stefanovic S., Krueger L., Olmstead R.G.;  
RT "Monophyly of the Convolvulaceae and circumscription of their major  
RT lineages based on DNA sequences of multiple chloroplast loci.";  
RL Am. J. Bot. 0:0-0(2002).

DR EMBL; AY100905; AAM55743.1; -.

KW Chloroplast.

FT NON\_TER 11 11

SQ SEQUENCE 11 AA; 1260 MW; 93736D59440861B1 CRC64;

Query Match 37.5%; Score 3; DB 8; Length 11;

Best Local Similarity 100.0%; Pred. No. 2.7e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GRI 7

|||

Db 6 GRI 8

#### RESULT 9

Q8MAZ3

ID Q8MAZ3 PRELIMINARY; PRT; 11 AA.

AC Q8MAZ3;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE PsbJ (Fragment).

GN PSBJ.

OS Maripa repens.

OG Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; lamiids; Solanales; Convolvulaceae; Maripa.

OX NCBI\_TaxID=197412;

RN [1]

RP SEQUENCE FROM N.A.

RA Stefanovic S., Krueger L., Olmstead R.G.;

RT "Monophyly of the Convolvulaceae and circumscription of their major

RT lineages based on DNA sequences of multiple chloroplast loci.";

RL Am. J. Bot. 0:0-0(2002).

DR EMBL; AY100936; AAM55865.1; -.

KW Chloroplast.

FT NON\_TER 11 11

SQ SEQUENCE 11 AA; 1260 MW; 93736D59440861B1 CRC64;

Query Match 37.5%; Score 3; DB 8; Length 11;

Best Local Similarity 100.0%; Pred. No. 2.7e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GRI 7

|||

Db 6 GRI 8

#### RESULT 10

Q8MBE1

ID Q8MBE1 PRELIMINARY; PRT; 11 AA.

AC Q8MBE1;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE PsbJ (Fragment).

GN PSBJ.  
 OS Ipomoea alba.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamiids; Solanales; Convolvulaceae; Ipomoea.  
 OX NCBI\_TaxID=89634;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stefanovic S., Krueger L., Olmstead R.G.;  
 RT "Monophyly of the Convolvulaceae and circumscription of their major  
 RT lineages based on DNA sequences of multiple chloroplast loci.";  
 RL Am. J. Bot. 0:0-0(2002).  
 DR EMBL; AY100861; AAM55568.1; -.  
 KW Chloroplast.  
 FT NON\_TER 11 11  
 SQ SEQUENCE 11 AA; 1260 MW; 93736D59440861B1 CRC64;

Query Match 37.5%; Score 3; DB 8; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GRI 7  
 |||  
 Db 6 GRI 8

# RESULT 11

Q8MB77

ID Q8MB77 PRELIMINARY; PRT; 11 AA.  
 AC Q8MB77;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE PsbJ (Fragment).  
 GN PSBJ.  
 OS Odonellia hirtiflora.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamiids; Solanales; Convolvulaceae; Odonellia.  
 OX NCBI\_TaxID=197424;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stefanovic S., Krueger L., Olmstead R.G.;  
 RT "Monophyly of the Convolvulaceae and circumscription of their major  
 RT lineages based on DNA sequences of multiple chloroplast loci.";  
 RL Am. J. Bot. 0:0-0(2002).  
 DR EMBL; AY100897; AAM55711.1; -.  
 KW Chloroplast.  
 FT NON\_TER 11 11  
 SQ SEQUENCE 11 AA; 1260 MW; 93736D59440861B1 CRC64;

Query Match 37.5%; Score 3; DB 8; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy            5 GRI 7  
               |||  
 Db            6 GRI 8

RESULT 12

Q8MB79

ID    Q8MB79            PRELIMINARY;            PRT;    11 AA.  
 AC    Q8MB79;  
 DT    01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT    01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT    01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE    PsbJ (Fragment).  
 GN    PSBJ.  
 OS    Aniseia argentina.  
 OG    Chloroplast.  
 OC    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC    Asteridae; lamiids; Solanales; Convolvulaceae; Aniseia.  
 OX    NCBI\_TaxID=197349;  
 RN    [1]  
 RP    SEQUENCE FROM N.A.  
 RA    Stefanovic S., Krueger L., Olmstead R.G.;  
 RT    "Monophyly of the Convolvulaceae and circumscription of their major  
 RT    lineages based on DNA sequences of multiple chloroplast loci.";  
 RL    Am. J. Bot. 0:0-0(2002).  
 DR    EMBL; AY100895; AAM55703.1; -.  
 KW    Chloroplast.  
 FT    NON\_TER            11        11  
 SQ    SEQUENCE    11 AA;    1260 MW;    93736D59440861B1 CRC64;

Query Match                    37.5%;    Score 3;    DB 8;    Length 11;  
 Best Local Similarity    100.0%;    Pred. No. 2.7e+03;  
 Matches        3;    Conservative        0;    Mismatches        0;    Indels        0;    Gaps        0;

Qy            5 GRI 7  
               |||  
 Db            6 GRI 8

RESULT 13

Q8MB97

ID    Q8MB97            PRELIMINARY;            PRT;    11 AA.  
 AC    Q8MB97;  
 DT    01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT    01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT    01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE    PsbJ (Fragment).  
 GN    PSBJ.  
 OS    Merremia peltata.  
 OG    Chloroplast.  
 OC    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC    Asteridae; lamiids; Solanales; Convolvulaceae; Merremia.  
 OX    NCBI\_TaxID=197416;  
 RN    [1]  
 RP    SEQUENCE FROM N.A.

RA Stefanovic S., Krueger L., Olmstead R.G.;  
 RT "Monophyly of the Convolvulaceae and circumscription of their major  
 RT lineages based on DNA sequences of multiple chloroplast loci.";   
 RL Am. J. Bot. 0:0-0(2002).  
 DR EMBL; AY100885; AAM55663.1; -.  
 KW Chloroplast.  
 FT NON\_TER 11 11  
 SQ SEQUENCE 11 AA; 1260 MW; 93736D59440861B1 CRC64;

Query Match 37.5%; Score 3; DB 8; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GRI 7  
 |||  
 Db 6 GRI 8

#### RESULT 14

Q31006

ID Q31006 PRELIMINARY; PRT; 12 AA.  
 AC Q31006;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Bota protein (Fragment).  
 GN BOTA.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93052564; PubMed=1428011;  
 RA Ellis S.A., Braem K.A., Morrison W.I.;  
 RT "Transmembrane and cytoplasmic domain sequences demonstrate at least  
 RT two expressed bovine MHC class I loci.";   
 RL Immunogenetics 37:49-56(1992).  
 DR EMBL; S47738; AAB23972.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 12 AA; 1306 MW; 6D9E2F805ABB5044 CRC64;

Query Match 37.5%; Score 3; DB 7; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GRI 7  
 |||  
 Db 4 GRI 6

#### RESULT 15

Q8MAX7

ID Q8MAX7 PRELIMINARY; PRT; 12 AA.  
 AC Q8MAX7;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE PsbJ (Fragment).  
 GN PSBJ.  
 OS Tridynamia megalantha.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; lamiids; Solanales; Convolvulaceae; Tridynamia.  
 OX NCBI\_TaxID=197451;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stefanovic S., Krueger L., Olmstead R.G.;  
 RT "Monophyly of the Convolvulaceae and circumscription of their major  
 RT lineages based on DNA sequences of multiple chloroplast loci.";  
 RL Am. J. Bot. 0:0-0(2002).  
 DR EMBL; AY100945; AAM55900.1; -.  
 KW Chloroplast.  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1316 MW; 92DAE36D59440861 CRC64;

Query Match 37.5%; Score 3; DB 8; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 GRI 7  
 |||  
 Db 6 GRI 8

Search completed: November 13, 2003, 10:38:17  
 Job time : 20.6667 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:53:11 ; Search time 25.6667 Seconds  
(without alignments)  
49.473 Million cell updates/sec

Title: US-09-228-866-9  
Perfect score: 8  
Sequence: 1 CTRITESC 8

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 358712

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

#### SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	8	100.0	8	18	AAW13420		Brain homing pepti
2	8	100.0	8	21	AAB07395		Brain homing pepti
3	8	100.0	8	22	AAE11801		Phage peptide #9 t
4	8	100.0	8	23	AAU10712		Brain homing pepti
5	5	62.5	7	22	AAM44549		H11 binding site c
6	5	62.5	13	9	AAP81811		Sequence of peptid
7	5	62.5	13	13	AAR22784		HIV-2 env protein
8	5	62.5	14	22	AAM00381		Human protein frag
9	5	62.5	14	22	AAM00382		Human protein frag
10	5	62.5	18	21	AAY51994		HIV-2 ROD isolate
11	4	50.0	7	16	AAR75853		Factor XI(a) plate
12	4	50.0	7	22	AAM45438		H11 binding site c
13	4	50.0	7	22	AAM45443		H11 binding site c
14	4	50.0	7	22	AAM45907		H11 binding site c
15	4	50.0	9	24	ABR07471		Human cancer-relat
16	4	50.0	9	24	ABR07476		Human cancer-relat
17	4	50.0	9	24	ABR07483		Human cancer-relat
18	4	50.0	9	24	ABR07498		Human cancer-relat
19	4	50.0	9	24	ABR07914		Human cancer-relat
20	4	50.0	9	24	ABR07938		Human cancer-relat
21	4	50.0	9	24	ABR07965		Human cancer-relat
22	4	50.0	9	24	ABR08114		Human cancer-relat
23	4	50.0	9	24	ABR08138		Human cancer-relat
24	4	50.0	9	24	ABR08165		Human cancer-relat
25	4	50.0	9	24	ABR08302		Human cancer-relat
26	4	50.0	9	24	ABR08318		Human cancer-relat
27	4	50.0	9	24	ABR08334		Human cancer-relat
28	4	50.0	9	24	ABR08745		Human cancer-relat
29	4	50.0	10	22	AAG87293		Saccharomyces cere
30	4	50.0	10	24	ABR07566		Human cancer-relat
31	4	50.0	10	24	ABR07571		Human cancer-relat
32	4	50.0	10	24	ABR07627		Human cancer-relat
33	4	50.0	10	24	ABR07798		Human cancer-relat
34	4	50.0	10	24	ABR07839		Human cancer-relat
35	4	50.0	10	24	ABR07992		Human cancer-relat
36	4	50.0	10	24	ABR07995		Human cancer-relat
37	4	50.0	10	24	ABR08213		Human cancer-relat
38	4	50.0	10	24	ABR08258		Human cancer-relat
39	4	50.0	10	24	ABR08260		Human cancer-relat
40	4	50.0	10	24	ABR08374		Human cancer-relat
41	4	50.0	10	24	ABR08428		Human cancer-relat
42	4	50.0	10	24	ABR08800		Human cancer-relat
43	4	50.0	10	24	ABR08900		Human cancer-relat
44	4	50.0	12	21	AAY95417		Anti-angiogenic pe
45	4	50.0	12	22	AAE23202		Human factor FHR-2

#### ALIGNMENTS

RESULT 1

AAW13420

ID AAW13420 standard; Peptide; 8 AA.

XX

AC AAW13420;

XX

DT 15-JAN-1998 (first entry)

XX

DE Brain homing peptide.

XX

KW Brain homing peptide; in vivo panning; screening; phage display;

KW drug delivery.

XX

OS Synthetic.

XX

PN WO9710507-A1.

XX

PD 20-MAR-1997.

XX

PF 10-SEP-1996; 96WO-US14600.

XX

PR 11-SEP-1995; 95US-0526710.

PR 11-SEP-1995; 95US-0526708.

XX

PA (LJOL-) LA JOLLA CANCER RES FOUND.

XX

PI Pasqualini R, Ruoslahti E;

XX

DR WPI; 1997-202359/18.

XX

PT Obtaining compound that homes to selected organ or tissue - by in

PT vivo panning method, specifically to identify brain, kidney,

PT angiogenic vasculature or tumour tissue homing peptide(s)

XX

PS Claim 15; Page 68; 75pp; English.

XX

CC This synthetic peptide is a claimed example of a brain-homing  
 CC peptide that was identified using a novel method for obtaining  
 CC molecules that home to a selected organ or tissue. This in vivo  
 CC panning method typically involves administering a phage display  
 CC library to a subject, and identifying expressed peptides which  
 CC home to the desired organ or tissue, e.g. brain, kidney, angiogenic  
 CC vascular tissue or tumour tissue. The isolated peptides (see  
 CC AAW13412-52, AAW11181-86) can be used to target e.g. drugs, toxins or  
 CC labels to the selected organ/tissue (claimed) or to identify and/or  
 CC isolate target molecules (claimed). The peptides can be directly  
 CC identified in vivo, as compared to prior art in vitro screening  
 CC methods, which require further examination to see if they maintain  
 CC specificity in vivo.

XX

SQ Sequence 8 AA;

Query Match 100.0%; Score 8; DB 18; Length 8;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTRITESC 8

Db                    |||||  
                      1 CTRITESC 8

RESULT 2

AAB07395

ID    AAB07395 standard; peptide; 8 AA.

XX

AC    AAB07395;

XX

DT    17-OCT-2000    (first entry)

XX

DE    Brain homing peptide # 9.

XX

KW    Brain; homing peptide; organ targeting; tissue targeting; mouse; cyclic.

XX

OS    Mus sp.

XX

FH    Key                    Location/Qualifiers

FT    Disulfide-bond    1..8

FT                                /note= "Can optionally form a cyclic peptide"

XX

PN    US6068829-A.

XX

PD    30-MAY-2000.

XX

PF    23-JUN-1997;    97US-0862855.

XX

PR    11-SEP-1995;    95US-0526710.

PR    10-MAR-1997;    97US-0813273.

XX

PA    (BURN-) BURNHAM INST.

XX

PI    Pasqualini R,   Ruoslahti E;

XX

DR    WPI; 2000-410850/35.

XX

PT    Identifying and recovering organ homing molecules or peptides by in  
PT    vivo panning comprises administering a library of diverse peptides  
PT    linked to a tag which facilitates recovery of these peptides -

XX

PS    Example 2; Column 17; 20pp; English.

XX

CC    The present sequence is a mouse brain homing peptide. This sequence was  
CC    identified by using in vivo panning to screen a library of potential  
CC    organ homing molecules. The present sequence can be used to direct a  
CC    moiety to a the brain tissue, by linking the moiety to the present  
CC    sequence. Examples of potential moieties are drugs, toxins or a  
CC    detectable label.

XX

SQ    Sequence    8 AA;

Query Match                    100.0%;    Score 8;    DB 21;    Length 8;

Best Local Similarity    100.0%;    Pred. No. 9.3e+05;

Matches    8;    Conservative    0;    Mismatches    0;    Indels    0;    Gaps    0;

Qy                    1 CTRITESC 8

Db

|||||||  
1 CTRITESC 8

RESULT 3

AAE11801

ID AAE11801 standard; peptide; 8 AA.

XX

AC AAE11801;

XX

DT 18-DEC-2001 (first entry)

XX

DE Phage peptide #9 targetted to brain.

XX

KW Enriched library fraction; brain; kidney; tumour; panning; diagnostic;

KW molecular medicine; drug delivery; peptidomimetic; pharmaceutical.

XX

OS Bacteriophage.

XX

PN US6296832-B1.

XX

PD 02-OCT-2001.

XX

PF 08-JAN-1999; 99US-0226985.

XX

PR 23-JUN-1997; 97US-0862855.

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Ruoslahti E, Pasqualini R;

XX

DR WPI; 2001-610691/70.

XX

PT Enriched library fraction comprising molecules recovered by in vivo

PT panning that selectively home to a selected organ or tissue useful for

PT treating disease or in diagnostic methods -

XX

PS Example 2; Column 17; 21pp; English.

XX

CC The invention relates to an enriched library fraction containing  
CC molecules that selectively home to a selected organ or tissue such as  
CC brain, kidney or tumour recovered by in vivo panning. The invention  
CC generally relates to the field of molecular medicine, drug delivery and  
CC to a method of invivo panning for identifying a molecule that homes to a  
CC specific organ. The molecules, e.g., peptides, peptidomimetics, proteins  
CC and fragments of proteins contained in an enriched library fraction may  
CC be administered to a subject as part of a pharmaceutical composition to  
CC treat disease or in diagnostic methods. The present sequence is a  
CC peptide from bacteriophage targetted to brain.

XX

SQ Sequence 8 AA;

Query Match 100.0%; Score 8; DB 22; Length 8;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTRITESC 8  
|||  
Db 1 CTRITESC 8

RESULT 4

AAU10712

ID AAU10712 standard; peptide; 8 AA.

XX

AC AAU10712;

XX

DT 12-MAR-2002 (first entry)

XX

DE Brain homing peptide #9 useful for delivery of target molecules.

XX

KW Organ targeting; tissue targeting; cancer; tumour homing molecule;

KW delivery of target molecule; brain homing peptide.

XX

OS Synthetic.

XX

PN US6306365-B1.

XX

PD 23-OCT-2001.

XX

PF 08-JAN-1999; 99US-0227906.

XX

PR 23-JUN-1997; 97US-0862855.

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Ruoslahti E, Pasqualini R;

XX

DR WPI; 2002-040196/05.

XX

PT Recovering molecules that home to an organ or tissue, useful for  
PT identifying molecules that home to a specific organ or tissue, e.g.  
PT identifying a tumour homing molecule to identify the presence of cancer,  
PT by in vivo panning of a library -

XX

PS Example 2; Column 17; 21pp; English.

XX

CC The present invention relates to a method of recovering molecules that  
CC home to a selected organ or tissue. The method comprises administering  
CC to the subject the library of diverse molecules, collecting a sample of  
CC the selected organ or tissue (e.g. brain or kidney), and recovering from  
CC the sample several molecules that home to the selected organ or tissue.  
CC The method is useful for identifying molecules, particularly useful for  
CC screening large number of molecules (e.g. peptides), that home to a  
CC specific organ. The identified molecule is useful for e.g. raising an  
CC antibody specific for a target molecule, targeting a desired moiety  
CC (e.g. drug, toxin or detectable label) to the selected organ.  
CC Specifically, the method is useful for identifying the presence of cancer  
CC in a subject by linking an appropriate moiety to a tumour homing  
CC molecule. The present method provides a direct means for identifying

CC molecules that specifically home to a selected organ and, therefore  
CC provides a significant advantage over previous methods, which require  
CC that a molecule identified using an in vitro screening method  
CC subsequently be examined to determine if it maintains its specificity in  
CC vivo. AAU10704-AAU10723 represent brain homing peptides described in  
CC the present invention.

XX

SQ Sequence 8 AA;

Query Match 100.0%; Score 8; DB 23; Length 8;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTRITESC 8  
|||  
Db 1 CTRITESC 8

#### RESULT 5

AAM44549

ID AAM44549 standard; Peptide; 7 AA.

XX

AC AAM44549;

XX

DT 25-OCT-2001 (first entry)

XX

DE H11 binding site consensus conforming peptide (CCP) #820.

XX

KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;  
KW immunogenically cross-reactive; cancer; immunogenic cancer cell;  
KW cytostatic; vaccine; tumour-specific immunogenic response inducer;  
KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;  
KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN CA2290722-A1.

XX

PD 08-JUN-2001.

XX

PF 08-DEC-1999; 99CA-2290722.

XX

PR 08-DEC-1999; 99CA-2290722.

XX

PA (NOVO-) NOVOPHARM BIOTECH INC.

XX

PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;

PI Entwistle JM, MacDonald GC;

XX

DR WPI; 2001-425937/46.

XX

PT Composition useful for treating and diagnosing cancer, comprises stress  
PT protein-peptide complexes associated with tumor, and isolated  
PT antigen-binding fragments of an antibody that binds specifically to the  
PT complex -

XX

PS Example 4; Page 103; 154pp; English.

XX

CC The present invention describes a composition (I) comprising stress  
CC protein-peptide complexes (SPPC) associated with tumours that is  
CC specifically immunogenically cross-reactive with cell surface-associated  
CC SPPCs specific to target cancer (TC). Also described is an isolated  
CC antigen-binding fragment of an antibody that binds specifically to SPPCs  
CC or a population of different SPPCs consisting of immunogenic cancer cell  
CC surface-associated SPPC of TC. (I) has cytostatic activity and can be  
CC used in vaccine production and as a tumour-specific immunogenic response  
CC inducer. (I) is useful for treating 71 types of cancers or tumours in a  
CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,  
CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural  
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including  
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting  
CC or imaging cancer cells, and to monitor the course of amelioration of  
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides  
CC which are used in the exemplification of the present invention.

XX

SQ Sequence 7 AA;

Query Match 62.5%; Score 5; DB 22; Length 7;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RITES 7

|||||

Db 3 RITES 7

#### RESULT 6

AAP81811

ID AAP81811 standard; protein; 13 AA.

XX

AC AAP81811;

XX

DT 25-MAR-2003 (updated)

DT 16-NOV-1990 (first entry)

XX

DE Sequence of peptide with immunological properties and structure of  
DE HIV-2 and SIV.1 glycoproteins.

XX

KW Antigen; immunogen; vaccine; diagnostic; AIDS.

XX

OS Immunodeficiency virus.

XX

PN WO8805440-A.

XX

PD 28-JUL-1988.

XX

PF 15-JAN-1988; 88WO-EP00025.

XX

PR 16-JAN-1987; 87US-0003764.

PR 11-FEB-1987; 87FR-0001739.

PR 15-APR-1987; 87FR-0005398.

XX

PA (INSP ) INST PASTEUR.



PA (ALIZ/) ALIZON M.  
PA (CNRS ) CENT NAT RECH SCI.  
XX  
PI Alizon M, Montagnier L, Guetard D, Clavfl F, Sonigo P, Guyader M;  
PI Tiollais P, Chakrabarti L, Desrosiers R;  
XX  
DR WPI; 1988-220290/31.  
XX  
PT New peptide(s) with immunological properties of HIV-2 envelope protein -  
PT have the structure of simian immune deficiency virus proteins,  
PT useful in diagnosis and of vaccine components  
XX  
PS Claim 12; Page 44; 86pp; French.  
XX  
CC New peptides which have immunological properties in common with those of  
CC the peptide skeleton of the envelope protein of HIV-2 and also have a  
CC peptide structure in common with that of SIV-1 glycoprotein are claimed.  
CC Antigenic and immunogenic conjugates contg. the peptides and a kit to  
CC detect HIV-2 in biological fluids are new. The peptides are useful for in  
CC vitro diagnosis of HIV-2 infection and some of them can be used as  
CC components of immunogens and vaccines against HIV. Antibodies raised  
CC against them can be used for treatment of AIDS.  
CC (Updated on 25-MAR-2003 to correct PF field.)  
CC (Updated on 25-MAR-2003 to correct PR field.)  
CC (Updated on 25-MAR-2003 to correct PA field.)  
XX  
SQ Sequence 13 AA;

Query Match 62.5%; Score 5; DB 9; Length 13;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITESC 8  
|||||  
Db 8 ITESC 12

RESULT 7  
AAR22784  
ID AAR22784 standard; Protein; 13 AA.  
XX  
AC AAR22784;  
XX  
DT 25-MAR-2003 (updated)  
DT 19-MAY-1992 (first entry)  
XX  
DE HIV-2 env protein specific immunological peptide 7.  
XX  
KW Human immunodeficiency virus; AIDS; envelope glycoprotein.  
XX  
OS Human immunodeficiency virus-2 ROD isolate.  
XX  
PN US5079342-A.  
XX  
PD 07-JAN-1992.  
XX  
PF 11-FEB-1987; 87US-0013477.

XX  
 PR 11-FEB-1987; 87US-0013477.  
 XX  
 PA (INSP ) INST PASTEUR.  
 XX  
 PI Alizon M, Montagnier L, Geutard D, Clavel F, Sonigo P, Guyader M;  
 XX  
 DR WPI; 1992-041067/05.  
 DR N-PSDB; AAQ20616.  
 XX  
 PT Peptide(s) corresp. to HIV-2 amino acid sequences - used in  
 PT diagnosis in vaccines and in prodn. of antibodies for diagnosis  
 XX  
 PS Claim 12; Page 30; 30pp; English.  
 XX  
 CC The amino acid sequence is that of a peptide which comprises the  
 CC immunological properties of the first portion of the envelope  
 CC glycoprotein of a ROD isolate HIV-2 virus. It can be used for  
 CC detecting HIV-2 infection and for producing antibodies for use in  
 CC diagnosis. It can also be used to induce protection against  
 CC infection by HIV-2 and in competitive assays to test the ability  
 CC of antiviral agents to prevent the virus from fixing on its target.  
 CC See also AAR20597-R20607 and AAR22771-R22793.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 SQ Sequence 13 AA;

Query Match 62.5%; Score 5; DB 13; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITESC 8  
 |||||  
 Db 8 ITESC 12

RESULT 8  
 AAM00381  
 ID AAM00381 standard; Peptide; 14 AA.  
 XX  
 AC AAM00381;  
 XX  
 DT 01-OCT-2001 (first entry)  
 XX  
 DE Human protein fragment SEQ ID NO: 929.  
 XX  
 KW Human; single nucleotide polymorphism; SNP; paternity test;  
 KW forensic test; aberrant protein expression.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200151670-A2.  
 XX  
 PD 19-JUL-2001.  
 XX  
 PF 05-JAN-2001; 2001WO-US00322.  
 XX

PR 07-JAN-2000; 2000US-0174962.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Shimkets RA, Leach MD;

XX

DR WPI; 2001-451871/48.

DR N-PSDB; AAH89498.

XX

PT Isolated human polynucleotides containing single nucleotide  
PT polymorphisms, useful for the treatment and diagnosis of e.g. cancer,  
PT infection and diabetes -

XX

PS Disclosure; Page 370; 475pp; English.

XX

CC The present invention relates to human nucleic acids containing single  
CC nucleotide polymorphisms (SNPs). These can be used in forensic and  
CC paternity tests, and to aid in the treatment of diseases associated with  
CC aberrant protein expression, including cancer, amyloidosis, diabetes,  
CC Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,  
CC glomerulonephritis, haemolytic anaemia, thrombocytopaenia, arthritis,  
CC meningitis, muscular disorders, dementia, neurological diseases, tuberous  
CC sclerosis, male infertility, hypercalcaemia, blood pressure disorders,  
CC osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or  
CC autoimmunity. The present sequence is a peptide encoded by a  
CC polymorphism-containing oligonucleotide fragment of the invention.

XX

SQ Sequence 14 AA;

Query Match 62.5%; Score 5; DB 22; Length 14;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITESC 8

|||||

Db 2 ITESC 6

RESULT 9

AAM00382

ID AAM00382 standard; Peptide; 14 AA.

XX

AC AAM00382;

XX

DT 01-OCT-2001 (first entry)

XX

DE Human protein fragment SEQ ID NO: 930.

XX

KW Human; single nucleotide polymorphism; SNP; paternity test;  
KW forensic test; aberrant protein expression.

XX

OS Homo sapiens.

XX

PN WO200151670-A2.

XX

PD 19-JUL-2001.

XX

PF 05-JAN-2001; 2001WO-US00322.

XX

PR 07-JAN-2000; 2000US-0174962.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Shimkets RA, Leach MD;

XX

DR WPI; 2001-451871/48.

DR N-PSDB; AAH89499.

XX

PT Isolated human polynucleotides containing single nucleotide  
PT polymorphisms, useful for the treatment and diagnosis of e.g. cancer,  
PT infection and diabetes -

XX

PS Disclosure; Page 371; 475pp; English.

XX

CC The present invention relates to human nucleic acids containing single  
CC nucleotide polymorphisms (SNPs). These can be used in forensic and  
CC paternity tests, and to aid in the treatment of diseases associated with  
CC aberrant protein expression, including cancer, amyloidosis, diabetes,  
CC Alzheimer's disease, Down's syndrome, oedema, lupus (SLE), vasculitis,  
CC glomerulonephritis, haemolytic anaemia, thrombocytopaenia, arthritis,  
CC meningitis, muscular disorders, dementia, neurological diseases, tuberculous  
CC sclerosis, male infertility, hypercalcaemia, blood pressure disorders,  
CC osteoporosis, pathogenic infections, hypercholesterolaemia, obesity or  
CC autoimmunity. The present sequence is a peptide encoded by a  
CC polymorphism-containing oligonucleotide fragment of the invention.

XX

SQ Sequence 14 AA;

Query Match 62.5%; Score 5; DB 22; Length 14;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITESC 8

|||||

Db 2 ITESC 6

#### RESULT 10

AAAY51994

ID AAY51994 standard; Protein; 18 AA.

XX

AC AAY51994;

XX

DT 11-JUL-2000 (first entry)

XX

DE HIV-2 ROD isolate antigenic peptide fragment env7.

XX

KW Diagnosis; probe; lymphocyte; virus; immunodiagnostic; infection;  
KW antiviral; ENV protein; antigenic.

XX

OS Human immunodeficiency virus type 2.

XX

PN US6054565-A.

XX

```

PD 25-APR-2000.
XX
PF 28-APR-1994; 94US-0234875.
XX
PR 11-FEB-1987; 87US-0013477.
PR 03-SEP-1991; 91US-0752368.
PR 20-DEC-1991; 91US-0810908.
PR 03-MAR-1986; 86US-0835228.
PR 06-OCT-1986; 86US-0916080.
PR 21-NOV-1986; 86US-0933184.
PR 16-JAN-1987; 87US-0003764.
XX
PA (INSP ) INST PASTEUR.
XX
PI Montagnier L, Clavel F, Guyader M, Geutard D, Sonigo P, Alizon M;
XX
DR WPI; 2000-328365/28.
DR N-PSDB; AAZ89635.
XX
PT Novel cloned nucleotide sequences homologous or identical to the
PT portion of genomic RNA of HIV-2 viruses useful as probes and in
PT diagnostic tests to diagnose HIV-2 infection -
XX
PS Example 6; Column 27-28; 33pp; English.
XX
CC This invention describes a novel cloned nucleic acid (I) of a human
CC immunodeficiency virus type 2 (HIV-2). (I) is capable of being used
CC as probes in diagnostic method to obtain the immunological reagents
CC necessary to diagnose an HIV-2 infection. These sequences may be used
CC as probes in hybridization reactions with the genetic material of
CC infected patients to indicate whether the RNA of the HIV-2 virus is
CC present in these patient's lymphocytes or whether an analogous DNA is
CC present. The genetic sequence of the HIV-2 virus may be used to create
CC the polypeptides encoded by these sequences. Specifically, these
CC polypeptides may be created by expression of the cDNA obtained from
CC bacterial, yeast or animal cells. These polypeptides may be used in
CC diagnostic tests such as immunofluorescence assays, radioimmunoassays
CC (RIA) and Western Blot tests. Monoclonal antibodies to these
CC polypeptides or fragments may be created and used in immunodiagnostic
CC tests. The polypeptides of the present invention may also be used as
CC immunogenic reagents to induce protection against infection by HIV-2
CC viruses. The polypeptides produced by recombinant-DNA techniques would
CC function as vaccine agents. The polypeptides may be used on competitive
CC assays to test the ability of various antiviral agents to determined
CC their ability to prevent the virus from fixing on its target.
CC AAY51988-Y51999 represent HIV-2 ROD isolate antigenic fragments encoded
CC by AAZ89629-Z89640 which are described in the method of the invention.
XX
SQ Sequence 18 AA;

Query Match 62.5%; Score 5; DB 21; Length 18;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITESC 8
   |||||
Db 8 ITESC 12

```

RESULT 11

AAR75853

ID AAR75853 standard; peptide; 7 AA.

XX

AC AAR75853;

XX

DT 11-MAR-1996 (first entry)

XX

DE Factor XI(a) platelet binding site peptide analogue.

XX

KW Factor XI(a); platelet binding site; peptide analogue; thrombosis;

KW antithrombotic agent; intrinsic coagulation.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Disulfide-bond 1..7

XX

PN WO9517420-A1.

XX

PD 29-JUN-1995.

XX

PF 02-DEC-1994; 94WO-US13885.

XX

PR 22-DEC-1993; 93US-0172002.

XX

PA (UYJE-) UNIV JEFFERSON THOMAS.

PA (UTEM ) UNIV TEMPLE.

XX

PI Baglia FA, Jameson BA, Walsh PN;

XX

DR WPI; 1995-240608/31.

XX

PT Peptide analogues of the factor XI platelet binding site - used to  
PT specifically inhibit coagulation reactions involving factor XI and  
PT factor XIa, for improved treatment of thrombosis.

XX

PS Claim 10; Page 78; 99pp; English.

XX

CC AAR75853 is a factor XI(a) platelet binding site peptide analogue,  
CC useful as an antithrombotic agent. The peptide specifically  
CC inhibits intrinsic coagulation reactions, while leaving extrinsic  
CC reactions intact. This permits normal hemostatic plug formation at  
CC a site of vascular injury, minimising the risk of bleeding during  
CC antithrombotic therapy.

XX

SQ Sequence 7 AA;

Query Match 50.0%; Score 4; DB 16; Length 7;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTRI 4

||||

Db 1 CTRI 4

RESULT 12

AAM45438

ID AAM45438 standard; Peptide; 7 AA.

XX

AC AAM45438;

XX

DT 25-OCT-2001 (first entry)

XX

DE H11 binding site consensus conforming peptide (CCP) #1709.

XX

KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;

KW immunogenically cross-reactive; cancer; immunogenic cancer cell;

KW cytostatic; vaccine; tumour-specific immunogenic response inducer;

KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;

KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN CA2290722-A1.

XX

PD 08-JUN-2001.

XX

PF 08-DEC-1999; 99CA-2290722.

XX

PR 08-DEC-1999; 99CA-2290722.

XX

PA (NOVO-) NOVOPHARM BIOTECH INC.

XX

PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;

PI Entwistle JM, MacDonald GC;

XX

DR WPI; 2001-425937/46.

XX

PT Composition useful for treating and diagnosing cancer, comprises stress

PT protein-peptide complexes associated with tumor, and isolated

PT antigen-binding fragments of an antibody that binds specifically to the

PT complex -

XX

PS Example 4; Page 107; 154pp; English.

XX

CC The present invention describes a composition (I) comprising stress

CC protein-peptide complexes (SPPC) associated with tumours that is

CC specifically immunogenically cross-reactive with cell surface-associated

CC SPPCs specific to target cancer (TC). Also described is an isolated

CC antigen-binding fragment of an antibody that binds specifically to SPPCs

CC or a population of different SPPCs consisting of immunogenic cancer cell

CC surface-associated SPPC of TC. (I) has cytostatic activity and can be

CC used in vaccine production and as a tumour-specific immunogenic response

CC inducer. (I) is useful for treating 71 types of cancers or tumours in a

CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,

CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural

CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including

CC vaccines. (I) is useful for diagnostic and palliative use, for detecting

CC or imaging cancer cells, and to monitor the course of amelioration of

CC malignancy in an individual. AAM43707 to AAM47109 represent peptides  
CC which are used in the exemplification of the present invention.

XX

SQ Sequence 7 AA;

Query Match 50.0%; Score 4; DB 22; Length 7;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RITE 6

||||

Db 3 RITE 6

RESULT 13

AAM45443

ID AAM45443 standard; Peptide; 7 AA.

XX

AC AAM45443;

XX

DT 25-OCT-2001 (first entry)

XX

DE H11 binding site consensus conforming peptide (CCP) #1714.

XX

KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;

KW immunogenically cross-reactive; cancer; immunogenic cancer cell;

KW cytostatic; vaccine; tumour-specific immunogenic response inducer;

KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;

KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN CA2290722-A1.

XX

PD 08-JUN-2001.

XX

PF 08-DEC-1999; 99CA-2290722.

XX

PR 08-DEC-1999; 99CA-2290722.

XX

PA (NOVO-) NOVOPHARM BIOTECH INC.

XX

PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;

PI Entwistle JM, MacDonald GC;

XX

DR WPI; 2001-425937/46.

XX

PT Composition useful for treating and diagnosing cancer, comprises stress

PT protein-peptide complexes associated with tumor, and isolated

PT antigen-binding fragments of an antibody that binds specifically to the

PT complex -

XX

PS Example 4; Page 107; 154pp; English.

XX

CC The present invention describes a composition (I) comprising stress  
CC protein-peptide complexes (SPPC) associated with tumours that is



CC specifically immunogenically cross-reactive with cell surface-associated  
CC SPPCs specific to target cancer (TC). Also described is an isolated  
CC antigen-binding fragment of an antibody that binds specifically to SPPCs  
CC or a population of different SPPCs consisting of immunogenic cancer cell  
CC surface-associated SPPC of TC. (I) has cytostatic activity and can be  
CC used in vaccine production and as a tumour-specific immunogenic response  
CC inducer. (I) is useful for treating 71 types of cancers or tumours in a  
CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,  
CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural  
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including  
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting  
CC or imaging cancer cells, and to monitor the course of amelioration of  
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides  
CC which are used in the exemplification of the present invention.

XX

SQ Sequence 7 AA;

Query Match 50.0%; Score 4; DB 22; Length 7;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RITE 6

||||

Db 3 RITE 6

#### RESULT 14

AAM45907

ID AAM45907 standard; Peptide; 7 AA.

XX

AC AAM45907;

XX

DT 25-OCT-2001 (first entry)

XX

DE H11 binding site consensus conforming peptide (CCP) #2178.

XX

KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;  
KW immunogenically cross-reactive; cancer; immunogenic cancer cell;  
KW cytostatic; vaccine; tumour-specific immunogenic response inducer;  
KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;  
KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN CA2290722-A1.

XX

PD 08-JUN-2001.

XX

PF 08-DEC-1999; 99CA-2290722.

XX

PR 08-DEC-1999; 99CA-2290722.

XX

PA (NOVO-) NOVOPHARM BIOTECH INC.

XX

PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;

PI Entwistle JM, MacDonald GC;

XX  
 DR WPI; 2001-425937/46.  
 XX  
 PT Composition useful for treating and diagnosing cancer, comprises stress  
 PT protein-peptide complexes associated with tumor, and isolated  
 PT antigen-binding fragments of an antibody that binds specifically to the  
 PT complex -  
 XX  
 PS Example 4; Page 108; 154pp; English.  
 XX  
 CC The present invention describes a composition (I) comprising stress  
 CC protein-peptide complexes (SPPC) associated with tumours that is  
 CC specifically immunogenically cross-reactive with cell surface-associated  
 CC SPPCs specific to target cancer (TC). Also described is an isolated  
 CC antigen-binding fragment of an antibody that binds specifically to SPPCs  
 CC or a population of different SPPCs consisting of immunogenic cancer cell  
 CC surface-associated SPPC of TC. (I) has cytostatic activity and can be  
 CC used in vaccine production and as a tumour-specific immunogenic response  
 CC inducer. (I) is useful for treating 71 types of cancers or tumours in a  
 CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,  
 CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural  
 CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including  
 CC vaccines. (I) is useful for diagnostic and palliative use, for detecting  
 CC or imaging cancer cells, and to monitor the course of amelioration of  
 CC malignancy in an individual. AAM43707 to AAM47109 represent peptides  
 CC which are used in the exemplification of the present invention.  
 XX  
 SQ Sequence 7 AA;

Query Match 50.0%; Score 4; DB 22; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RITE 6  
 ||||  
 Db 3 RITE 6

# RESULT 15

ABR07471

ID ABR07471 standard; Peptide; 9 AA.

XX

AC ABR07471;

XX

DT 19-MAY-2003 (first entry)

XX

DE Human cancer-related protein 151P1C7A HLA peptide #6.

XX

KW Human; cytostatic; vaccine; cancer; immune response; HLA;  
 KW human leukocyte antigen.

XX

OS Homo sapiens.

XX

PN WO200283921-A2.

XX

PD 24-OCT-2002.

XX

PF 10-APR-2002; 2002WO-US11654.

XX

PR 10-APR-2001; 2001US-282739P.

PR 10-APR-2001; 2001US-283112P.

PR 25-APR-2001; 2001US-286630P.

XX

PA (AGEN-) AGENSYS INC.

XX

PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;

PI Morrison K, Morrison RK, Raitano AB;

XX

DR WPI; 2003-075555/07.

XX

PT New composition comprising a substance that modulates the structure of  
PT proteins and polynucleotides, useful for therapeutic, prognostic and  
PT diagnostic reagents for eliciting cellular or humoral immune response  
PT in cancer patients -

XX

PS Claim 13; Page 187; 1021pp; English.

XX

CC The present invention relates to novel human cancer-related genes and  
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and  
CC proteins are useful for eliciting a humoral or cellular immune response.  
CC The genes are useful as probes and primers for the amplification and/or  
CC detection of genes, mRNAs or their fragments, as reagents for the  
CC diagnosis and/or prognosis of cancer, as coding sequences capable of  
CC directing the expression of the protein, as tools for modulating or  
CC inhibiting the expression of genes and/or translation of transcripts, and  
CC as therapeutic agents. The proteins and peptides are useful as  
CC therapeutic, prognostic and diagnostic reagents for cancer. The present  
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example  
CC from the invention.

XX

SQ Sequence 9 AA;

Query Match 50.0%; Score 4; DB 24; Length 9;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITES 7

||||

Db 1 ITES 4

Search completed: November 13, 2003, 10:32:56

Job time : 25.6667 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 10:38:27 ; Search time 15.6667 Seconds  
(without alignments)

93.222 Million cell updates/sec

Title: US-09-228-866-9  
Perfect score: 8  
Sequence: 1 CTRITESC 8

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 666188 seqs, 182559486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 124183

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	4	50.0	12	11	US-09-805-337A-34	Sequence 34, Appl
2	4	50.0	14	12	US-10-005-549-28	Sequence 28, Appl
3	4	50.0	15	10	US-09-867-852-38	Sequence 38, Appl
4	4	50.0	16	9	US-09-107-058-10	Sequence 10, Appl
5	4	50.0	16	10	US-09-945-825-16	Sequence 16, Appl
6	4	50.0	16	12	US-10-161-791-203	Sequence 203, App
7	4	50.0	19	12	US-10-029-386-28088	Sequence 28088, A
8	4	50.0	21	11	US-09-774-639-367	Sequence 367, App
9	4	50.0	21	11	US-09-969-730-245	Sequence 245, App
10	3	37.5	7	10	US-09-867-852-90	Sequence 90, Appl

11	3	37.5	7	10	US-09-813-718-33	Sequence 33, Appl
12	3	37.5	7	11	US-09-995-973-20	Sequence 20, Appl
13	3	37.5	7	11	US-09-995-973-21	Sequence 21, Appl
14	3	37.5	7	11	US-09-996-484-20	Sequence 20, Appl
15	3	37.5	7	11	US-09-996-484-21	Sequence 21, Appl
16	3	37.5	7	12	US-10-052-578-271	Sequence 271, App
17	3	37.5	7	12	US-10-253-532-109	Sequence 109, App
18	3	37.5	7	12	US-10-253-532-115	Sequence 115, App
19	3	37.5	7	12	US-10-053-520-271	Sequence 271, App
20	3	37.5	7	12	US-10-053-498B-271	Sequence 271, App
21	3	37.5	7	14	US-10-066-151-9	Sequence 9, Appli
22	3	37.5	7	15	US-10-155-922-122	Sequence 122, App
23	3	37.5	7	15	US-10-155-922-127	Sequence 127, App
24	3	37.5	8	9	US-09-765-086-189	Sequence 189, App
25	3	37.5	8	10	US-09-910-552-10	Sequence 10, Appl
26	3	37.5	8	12	US-10-052-578-30	Sequence 30, Appl
27	3	37.5	8	12	US-10-253-532-96	Sequence 96, Appl
28	3	37.5	8	12	US-10-253-532-103	Sequence 103, App
29	3	37.5	8	12	US-10-253-532-108	Sequence 108, App
30	3	37.5	8	12	US-10-348-504-46	Sequence 46, Appl
31	3	37.5	8	12	US-10-348-504-47	Sequence 47, Appl
32	3	37.5	8	12	US-10-348-504-48	Sequence 48, Appl
33	3	37.5	8	12	US-10-348-504-49	Sequence 49, Appl
34	3	37.5	8	12	US-10-348-504-54	Sequence 54, Appl
35	3	37.5	8	12	US-10-348-504-55	Sequence 55, Appl
36	3	37.5	8	12	US-10-348-504-56	Sequence 56, Appl
37	3	37.5	8	12	US-10-348-504-57	Sequence 57, Appl
38	3	37.5	8	12	US-10-348-504-62	Sequence 62, Appl
39	3	37.5	8	12	US-10-348-504-63	Sequence 63, Appl
40	3	37.5	8	12	US-10-348-504-64	Sequence 64, Appl
41	3	37.5	8	12	US-10-348-504-65	Sequence 65, Appl
42	3	37.5	8	12	US-10-348-504-75	Sequence 75, Appl
43	3	37.5	8	12	US-10-348-504-80	Sequence 80, Appl
44	3	37.5	8	12	US-10-348-504-81	Sequence 81, Appl
45	3	37.5	8	12	US-10-348-504-82	Sequence 82, Appl

#### ALIGNMENTS

```

RESULT 1
US-09-805-337A-34
; Sequence 34, Application US/09805337A
; Publication No. US20030049831A1
; GENERAL INFORMATION:
; APPLICANT: Baxter Healthcare Corporation
; TITLE OF INVENTION: A NOVEL FACTOR-H RELATED PROTEIN 5 AND ANTIBODIES THERETO
; FILE REFERENCE: DI-5585L US (BXTD 9000.1)
; CURRENT APPLICATION NUMBER: US/09/805,337A
; CURRENT FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/188,670
; PRIOR FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 12
; TYPE: PRT

```

; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Homology with human FHR-2 protein amino acids 47-58  
US-09-805-337A-34

Query Match 50.0%; Score 4; DB 11; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TRIT 5  
|||  
Db 4 TRIT 7

RESULT 2

US-10-005-549-28

; Sequence 28, Application US/10005549  
; Publication No. US20030190613A1  
; GENERAL INFORMATION:  
; APPLICANT: BOWEN, MICHAEL A.  
; APPLICANT: WU, YULI  
; APPLICANT: YANG, WEN-PIN  
; APPLICANT: FINGER, JOSHUA  
; APPLICANT: NADLER, STEVEN  
; APPLICANT: CARROLL, PAMELA  
; APPLICANT: FEDER, JOHN  
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING AN ACTIVATED HUMAN  
; TITLE OF INVENTION: T-LYMPHOCYTE-DERIVED PROTEIN RELATED TO UBIQUITIN  
; TITLE OF INVENTION: CONJUGATING ENZYME  
; FILE REFERENCE: D0034np  
; CURRENT APPLICATION NUMBER: US/10/005,549  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/308,706  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/244,688  
; PRIOR FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 28  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-005-549-28

Query Match 50.0%; Score 4; DB 12; Length 14;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITES 7  
|||  
Db 7 ITES 10

RESULT 3

US-09-867-852-38

; Sequence 38, Application US/09867852  
; Patent No. US20020147324A1

```

; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Staskawicz, Brian J.
; APPLICANT: Brent, Andrew F.
; APPLICANT: Dahlbeck, Douglas
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kunkel, Barbara N.
; APPLICANT: Mindrinos, Michael N.
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND
; TITLE OF INVENTION: DETECTION METHODS
; FILE REFERENCE: 00786/254002
; CURRENT APPLICATION NUMBER: US/09/867,852
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/301,085
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/310,912
; PRIOR FILING DATE: EARLIER FILING DATE: 1994-09-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/227,360
; PRIOR FILING DATE: EARLIER FILING DATE: 1994-04-13
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-867-852-38

```

```

Query Match          50.0%; Score 4; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      4 ITES 7
        ||||
Db      1 ITES 4

```

#### RESULT 4

```

US-09-107-058-10
; Sequence 10, Application US/09107058
; Patent No. US20010010922A1
; GENERAL INFORMATION:
; APPLICANT: Dalla-Favera, Riccardo
; APPLICANT: Niu, Hui-Feng
; TITLE OF INVENTION: CLONING AND USES OF THE GENETIC
; TITLE OF INVENTION: LOCUS bcl-6
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/09/107,058
;   FILING DATE:
;   CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;   NAME:  White, John P.
;   REGISTRATION NUMBER:  28,678
;   REFERENCE/DOCKET NUMBER:  0575/43771-A-PCT-US
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (212) 278-0400
;   TELEFAX:  (212) 391-0525
;   TELEX:  422523 COOP UI
;   INFORMATION FOR SEQ ID NO:  10:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:  16 amino acids
;   TYPE:  amino acid
;   STRANDEDNESS:  single
;   TOPOLOGY:  linear
;   MOLECULE TYPE:  protein
US-09-107-058-10

```

```

Query Match          50.0%;  Score 4;  DB 9;  Length 16;
Best Local Similarity 100.0%;  Pred. No. 2.4e+02;
Matches      4;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy      5  TESC 8
        ||||
Db     10  TESC 13

```

```

RESULT 5
US-09-945-825-16
; Sequence 16, Application US/09945825
; Patent No. US20020106669A1
; GENERAL INFORMATION:
; APPLICANT: NOBUHITO, SONE
; TITLE OF INVENTION: Respiratory Chain Enzyme Genes of Coryneform Bacteria
; FILE REFERENCE: 213639US-8222-10-0
; CURRENT APPLICATION NUMBER: US/09/945,825
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: JP 2000-270283
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-945-825-16

```

```

Query Match          50.0%;  Score 4;  DB 10;  Length 16;
Best Local Similarity 100.0%;  Pred. No. 2.4e+02;
Matches      4;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy      3  RITE 6

```



Db

||||  
8 RITE 11

RESULT 6

US-10-161-791-203

; Sequence 203, Application US/10161791  
; Publication No. US20030186863A1  
; GENERAL INFORMATION:  
; APPLICANT: SPARKS, Andrew B.  
; APPLICANT: KAY, Brian K.  
; APPLICANT: THORN, Judith M.  
; APPLICANT: QUILLIAM, Lawrence A.  
; APPLICANT: DER, Channing J.  
; APPLICANT: FOWLKES, Dana M.  
; APPLICANT: RIDER, James E.  
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF  
; TITLE OF INVENTION: ISOLATING AND USING SAME  
; NUMBER OF SEQUENCES: 467  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/161,791  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/602,999  
; FILING DATE: 16-FEB-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Misrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-202  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 203:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-10-161-791-203

Query Match 50.0%; Score 4; DB 12; Length 16;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy            2 TRIT 5  
              ||||  
Db            13 TRIT 16

RESULT 7

US-10-029-386-28088

; Sequence 28088, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES  
USEFUL FOR GENE  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 28088  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL158832.2  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5  
US-10-029-386-28088

Query Match                    50.0%; Score 4; DB 12; Length 19;  
Best Local Similarity    100.0%; Pred. No. 2.8e+02;  
Matches        4; Conservative        0; Mismatches        0; Indels        0; Gaps        0;

Qy            1 CTRI 4  
              ||||  
Db            1 CTRI 4

RESULT 8

US-09-774-639-367

; Sequence 367, Application US/09774639  
; Publication No. US20030003555A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 90 Human Secreted Proteins  
; FILE REFERENCE: PZ013P1  
; CURRENT APPLICATION NUMBER: US/09/774,639  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/244,112

; PRIOR FILING DATE: EARLIER FILING DATE: 1999-02-04  
; NUMBER OF SEQ ID NOS: 371  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 367  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-774-639-367

Query Match 50.0%; Score 4; DB 11; Length 21;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TRIT 5  
|||  
Db 7 TRIT 10

RESULT 9

US-09-969-730-245  
; Sequence 245, Application US/09969730  
; Publication No. US20030054443A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 90 Human Secreted Proteins  
; FILE REFERENCE: PZ013P2  
; CURRENT APPLICATION NUMBER: US/09/969,730  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: 09/774,639  
; PRIOR FILING DATE: 2001-02-01  
; PRIOR APPLICATION NUMBER: 60/238,291  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 09/244,112  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: PCT/US98/16235  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: 60/056,371  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,732  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,366  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,364  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,370  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,367  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,365  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,731  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,557  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/056,563  
; PRIOR FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 60/055,970

```

; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,986
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,311
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,808
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,803
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,809
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,806
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,310
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,798
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,309
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,312
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/054,807
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: 60/055,386
; PRIOR FILING DATE: 1997-08-05
; NUMBER OF SEQ ID NOS: 373
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 245
;   LENGTH: 21
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-969-730-245

```

```

Query Match          50.0%;  Score 4;  DB 11;  Length 21;
Best Local Similarity 100.0%;  Pred. No. 3.1e+02;
Matches      4;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy      2 TRIT 5
        ||||
Db      7 TRIT 10

```

```

RESULT 10
US-09-867-852-90
; Sequence 90, Application US/09867852
; Patent No. US20020147324A1
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Staskawicz, Brian J.
; APPLICANT: Brent, Andrew F.
; APPLICANT: Dahlbeck, Douglas
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kunkel, Barbara N.
; APPLICANT: Mindrinos, Michael N.
; APPLICANT: Yu, Guo-Liang

```

```

; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND
; TITLE OF INVENTION: DETECTION METHODS
; FILE REFERENCE: 00786/254002
; CURRENT APPLICATION NUMBER: US/09/867,852
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/301,085
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/310,912
; PRIOR FILING DATE: EARLIER FILING DATE: 1994-09-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/227,360
; PRIOR FILING DATE: EARLIER FILING DATE: 1994-04-13
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 90
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-867-852-90

```

```

Query Match          37.5%; Score 3; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches      3; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy          2 TRI 4
             |||
Db          4 TRI 6

```

# RESULT 11

```

US-09-813-718-33
; Sequence 33, Application US/09813718
; Publication No. US20020182666A1
; GENERAL INFORMATION:
; APPLICANT: Schimmel, Paul
; APPLICANT: Wakasugi, Keisuke
; TITLE OF INVENTION: Human Aminoacyl-tRNA Synthetase Polypeptides Useful For
; TITLE OF INVENTION: The Regulation of Angiogenesis
; FILE REFERENCE: 00-221
; CURRENT APPLICATION NUMBER: US/09/813,718
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Calcareo sp.
US-09-813-718-33

```

```

Query Match          37.5%; Score 3; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches      3; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy          3 RIT 5
             |||
Db          4 RIT 6

```

RESULT 12

US-09-995-973-20

```
; Sequence 20, Application US/09995973
; Publication No. US20030024006A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: ULLMAN, Christopher G.
; TITLE OF INVENTION: GENE SWITCHES
; FILE REFERENCE: 8325-2003 / G7-US1
; CURRENT APPLICATION NUMBER: US/09/995,973
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: zinc finger
; OTHER INFORMATION: binding domain
```

US-09-995-973-20

```
Query Match          37.5%; Score 3; DB 11; Length 7;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches      3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2 TRI 4
        |||
Db      4 TRI 6
```

RESULT 13

US-09-995-973-21

```
; Sequence 21, Application US/09995973
; Publication No. US20030024006A1
; GENERAL INFORMATION:
; APPLICANT: CHOO, Yen
; APPLICANT: ULLMAN, Christopher G.
; TITLE OF INVENTION: GENE SWITCHES
; FILE REFERENCE: 8325-2003 / G7-US1
; CURRENT APPLICATION NUMBER: US/09/995,973
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: zinc finger
; OTHER INFORMATION: binding domain
```

US-09-995-973-21

```
Query Match          37.5%; Score 3; DB 11; Length 7;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches      3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy            2 TRI 4  
             |||  
Db            4 TRI 6

RESULT 14

US-09-996-484-20

; Sequence 20, Application US/09996484  
; Publication No. US20030092010A1  
; GENERAL INFORMATION:  
; APPLICANT: CHOO, Yen  
; APPLICANT: ULLMAN, Christopher G.  
; TITLE OF INVENTION: MOLECULAR SWITCHES  
; FILE REFERENCE: 8325-2004 / G8-US1  
; CURRENT APPLICATION NUMBER: US/09/996,484  
; CURRENT FILING DATE: 2002-04-08  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 20  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: zinc finger  
; OTHER INFORMATION: binding domain  
US-09-996-484-20

Query Match                    37.5%; Score 3; DB 11; Length 7;  
Best Local Similarity    100.0%; Pred. No. 6e+05;  
Matches       3; Conservative    0; Mismatches    0; Indels       0; Gaps       0;

Qy            2 TRI 4  
             |||  
Db            4 TRI 6

RESULT 15

US-09-996-484-21

; Sequence 21, Application US/09996484  
; Publication No. US20030092010A1  
; GENERAL INFORMATION:  
; APPLICANT: CHOO, Yen  
; APPLICANT: ULLMAN, Christopher G.  
; TITLE OF INVENTION: MOLECULAR SWITCHES  
; FILE REFERENCE: 8325-2004 / G8-US1  
; CURRENT APPLICATION NUMBER: US/09/996,484  
; CURRENT FILING DATE: 2002-04-08  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 21  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: zinc finger  
; OTHER INFORMATION: binding domain  
US-09-996-484-21

Query Match 37.5%; Score 3; DB 11; Length 7;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TRI 4  
|||  
Db 4 TRI 6

Search completed: November 13, 2003, 11:12:33  
Job time : 15.6667 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 10:26:01 ; Search time 7.83333 Seconds  
(without alignments)  
98.215 Million cell updates/sec

Title: US-09-228-866-9  
Perfect score: 8  
Sequence: 1 CTRITESC 8

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3752

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : PIR\_76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description
-----	1	4	50.0	18	2	S58277	insulin-like growt



2	3	37.5	7	1	XEYDGD	galactose oxidase
3	3	37.5	10	2	S65728	hemoglobin, extrac
4	3	37.5	11	2	PC4267	ribosomal protein
5	3	37.5	12	2	I39390	acetylcholine rece
6	3	37.5	13	2	PQ0491	self-incompatibili
7	3	37.5	13	2	D56661	S-locus specific g
8	3	37.5	13	2	PT0305	Ig heavy chain CRD
9	3	37.5	13	2	D61458	Ig kappa chain V-I
10	3	37.5	13	2	E61458	Ig kappa chain V-I
11	3	37.5	15	2	PA0062	fumarate hydratase
12	3	37.5	15	2	A36527	juvenile-hormone e
13	3	37.5	15	2	A49480	major immunophilin
14	3	37.5	15	2	B45115	peptidylprolyl iso
15	3	37.5	17	2	S05671	hirudin Ia - medic
16	3	37.5	18	2	B45138	arsenite oxidase I
17	3	37.5	18	2	C59089	theta defensin-1 -
18	3	37.5	19	2	PS0236	trypsin inhibitor
19	3	37.5	19	2	C56661	S-locus specific g
20	3	37.5	19	2	PQ0492	self-incompatibili
21	3	37.5	19	2	PH0793	T-cell receptor al
22	3	37.5	19	2	S11611	ribosomal protein
23	3	37.5	20	2	A57106	hull allergen Gly
24	3	37.5	20	2	A60801	acrosome stabilizi
25	2	25.0	7	2	S19630	ribosomal protein
26	2	25.0	7	2	S16365	opacity protein P.
27	2	25.0	7	2	PT0087	ribulose-bisphosph
28	2	25.0	7	2	B34818	vicilin 57K chain
29	2	25.0	7	4	I55382	hypothetical pepti
30	2	25.0	8	2	A32523	peptidyl-dipeptida
31	2	25.0	8	2	S20162	leghemoglobin III
32	2	25.0	8	2	S59622	metallothionein is
33	2	25.0	8	2	PL0184	capsid protein VP-
34	2	25.0	8	2	S70727	ipgF protein - Shi
35	2	25.0	8	2	S63493	dissimilatory sulf
36	2	25.0	8	2	S71919	alcohol dehydrogen
37	2	25.0	8	2	PH0803	T-cell receptor al
38	2	25.0	8	2	I57018	gene Cftr protein
39	2	25.0	8	2	A61597	cytochrome P450 AL
40	2	25.0	8	2	A47618	beta-galactosidase
41	2	25.0	8	2	A25836	L-serine ammonia-l
42	2	25.0	8	2	T48890	hypothetical prote
43	2	25.0	9	2	C24180	fibrinogen beta ch
44	2	25.0	9	2	A29477	diuretic neuropept
45	2	25.0	9	2	PT0299	Ig heavy chain CRD

#### ALIGNMENTS

##### RESULT 1

S58277

insulin-like growth factor receptor type II - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 05-Nov-1999

C;Accession: S58277

R;Smrzka, O.W.; Stoger, R.; Kurzbauer, R.; Fae, I.; Fischer, G.F.; Barlow, D.P.  
submitted to the EMBL Data Library, January 1995

A;Description: Conservation of a methylation imprint and a putative imprinting box at the human IGF2R locus.  
A;Reference number: S58277  
A;Accession: S58277  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-18 <SMR>  
A;Cross-references: EMBL:X83702; NID:g929644; PIDN:CAA58675.1; PID:g929645  
C;Keywords: growth factor receptor

Query Match 50.0%; Score 4; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITES 7  
|||  
Db 4 ITES 7

#### RESULT 2

##### XEYDGD

galactose oxidase inhibitor - fungus (Cladobotryum dendroides)

C;Species: Cladobotryum dendroides

C;Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 31-Dec-1993

C;Accession: A01341

R;Avigad, G.; Markus, Z.

Fed. Proc. 31, 447, 1972

A;Reference number: A01341

A;Accession: A01341

A;Molecule type: protein

A;Residues: 1-7 <AVI>

C;Comment: The mycelia of this imperfect fungus produce the metalloenzyme galactose oxidase and its peptide inhibitor. The inhibitor, which can bind one copper ion per molecule and does not bind the apoenzyme, may inactivate the enzyme by binding to its prosthetic copper group.

C;Superfamily: galactose oxidase inhibitor

C;Keywords: copper

Query Match 37.5%; Score 3; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TES 7  
|||  
Db 5 TES 7

#### RESULT 3

##### S65728

hemoglobin, extracellular, chain d1 - earthworm (Lumbricus terrestris) (fragment)

C;Species: Lumbricus terrestris (common earthworm)

C;Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 13-Mar-1997

C;Accession: S65728

R;Fushitani, K.; Higashiyama, K.; Asao, M.; Hosokawa, K.

Biochim. Biophys. Acta 1292, 273-280, 1996

A;Title: Characterization of the constituent polypeptides of the extracellular hemoglobin from *Lumbricus terrestris*: heterogeneity and discovery of a new linker chain L4.

A;Reference number: S65721; MUID:96176855; PMID:8597573

A;Accession: S65728

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-10 <FUS>

Query Match 37.5%; Score 3; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TES 7  
|||  
Db 5 TES 7

#### RESULT 4

PC4267

ribosomal protein L12.1 - rice (fragment)

C;Species: *Oryza sativa* (rice)

C;Date: 28-May-1997 #sequence\_revision 18-Jul-1997 #text\_change 18-Jul-1997

C;Accession: PC4267

R;Kawakami, T.; Kamo, M.; Chen, M.C.; Tsugita, A.

submitted to JIPID, April 1997

A;Reference number: PC4267

A;Accession: PC4267

A;Molecule type: protein

A;Residues: 1-11 <KAW>

A;Experimental source: strain Japonica Nihonbare

Query Match 37.5%; Score 3; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TES 7  
|||  
Db 4 TES 6

#### RESULT 5

I39390

acetylcholine receptor (alternative exon 5b) - human (fragment)

C;Species: *Homo sapiens* (man)

C;Date: 16-Feb-1996 #sequence\_revision 16-Feb-1996 #text\_change 21-Jul-2000

C;Accession: I39390

R;Mihovilovic, M.; Mai, Y.; Herbstreith, M.; Rubboli, F.; Tarroni, P.; Clementi, F.; Roses, A.D.

Biochem. Biophys. Res. Commun. 197, 137-144, 1993

A;Title: Splicing of an anti-sense Alu sequence generates a coding sequence variant for the alpha-3 subunit of a neuronal acetylcholine receptor.

A;Reference number: I39390; MUID:94071933; PMID:8250918

A;Accession: I39390

A;Status: preliminary; translation not shown

A;Molecule type: mRNA

A;Residues: 1-12 <RES>

A;Cross-references: GB:L18973; NID:g441143; PIDN:AAA86792.1; PID:g441144  
C;Keywords: alternative splicing; neurotransmitter receptor

Query Match 37.5%; Score 3; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ESC 8  
|||  
Db 2 ESC 4

#### RESULT 6

PQ0491

self-incompatibility locus glycoprotein delta - wild cabbage (fragment)

C;Species: Brassica oleracea (wild cabbage)

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 31-Oct-1997

C;Accession: PQ0491

R;Gaude, T.; Friry, A.; Heizmann, P.; Mariac, C.; Rougier, M.; Fobis, I.; Dumas, C.

Plant Cell 5, 75-86, 1993

A;Title: Expression of a self-incompatibility gene in a self-compatible line of Brassica oleracea.

A;Reference number: JQ1733; MUID:93177215; PMID:8439745

A;Accession: PQ0491

A;Molecule type: protein

A;Residues: 1-13 <GAU>

A;Experimental source: stigma, var. acephala P57Si

C;Superfamily: S-receptor kinase; protein kinase homology; S-locus-specific glycoprotein homology

C;Keywords: glycoprotein

Query Match 37.5%; Score 3; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TES 7  
|||  
Db 7 TES 9

#### RESULT 7

D56661

S-locus specific glycoprotein (allele S3) - wild cabbage (fragment)

C;Species: Brassica oleracea (wild cabbage)

C;Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 05-Jan-1996

C;Accession: D56661

R;Gaude, T.; Denoroy, L.; Dumas, C.

Electrophoresis 12, 646-653, 1991

A;Title: Use of a fast protein electrophoretic purification procedure for N-terminal sequence analysis to identify S-locus related proteins in stigmas of Brassica oleracea.

A;Reference number: A56661; MUID:92090397; PMID:1752245

A;Accession: D56661

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-13 <GAU>

A;Experimental source: stigma extracts, var. acephala  
A;Note: sequence extracted from NCBI backbone (NCBIP:72300)  
C;Comment: This glycoprotein, expressed only in stigmas, plays an important role in the prevention of self-fertilization.  
C;Keywords: glycoprotein; polymorphism

Query Match 37.5%; Score 3; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TES 7  
|||  
Db 7 TES 9

#### RESULT 8

PT0305

Ig heavy chain CRD3 region (clone 5-121) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996

C;Accession: PT0305

R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
J. Exp. Med. 173, 395-407, 1991

A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and joining segments in adult human peripheral blood B lymphocytes.

A;Reference number: PT0222; MUID:91108337; PMID:1899102

A;Accession: PT0305

A;Molecule type: DNA

A;Residues: 1-13 <YAM>

A;Experimental source: B lymphocyte

C;Keywords: heterotetramer; immunoglobulin

Query Match 37.5%; Score 3; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTR 3  
|||  
Db 4 CTR 6

#### RESULT 9

D61458

Ig kappa chain V-IV region (DEP) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 15-Oct-1994 #sequence\_revision 15-Oct-1994 #text\_change 16-Aug-1996

C;Accession: D61458; PL0158

R;Brouet, J.C.; Dellagi, K.; Gendron, M.C.; Chevalier, A.; Schmitt, C.;  
Mihaesco, E.

J. Exp. Med. 170, 1551-1558, 1989

A;Title: Expression of a public idiotype by human monoclonal IgM directed to myelin-associated glycoprotein and characterization of the variability subgroup of their heavy and light chains.

A;Reference number: A61458; MUID:90039128; PMID:2478651

A;Accession: D61458

A;Molecule type: protein

A;Residues: 1-13 <BRO>

C;Comment: This protein is one of monoclonal IgM reactive with myelin-associated glycoprotein.

C;Keywords: heterotetramer; immunoglobulin

Query Match 37.5%; Score 3; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TES 7  
|||  
Db 5 TES 7

#### RESULT 10

E61458

Ig kappa chain V-IV region (FUE) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 15-Oct-1994 #sequence\_revision 15-Oct-1994 #text\_change 16-Aug-1996

C;Accession: E61458

R;Brouet, J.C.; Dellagi, K.; Gendron, M.C.; Chevalier, A.; Schmitt, C.; Mihaesco, E.

J. Exp. Med. 170, 1551-1558, 1989

A;Title: Expression of a public idiotype by human monoclonal IgM directed to myelin-associated glycoprotein and characterization of the variability subgroup of their heavy and light chains.

A;Reference number: A61458; MUID:90039128; PMID:2478651

A;Accession: E61458

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-13 <BRO>

C;Keywords: heterotetramer; immunoglobulin

Query Match 37.5%; Score 3; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TES 7  
|||  
Db 5 TES 7

#### RESULT 11

PA0062

fumarate hydratase (EC 4.2.1.2) - fungus (Fusarium sporotrichioides) (fragment)

C;Species: Fusarium sporotrichioides

C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Mar-2001

C;Accession: PA0062

R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A. submitted to JIPID, October 1994

A;Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporotrichioides proteins.

A;Reference number: PA0051

A;Accession: PA0062

A;Molecule type: protein

A;Residues: 1-15 <CHO>

C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 37.5%; Score 3; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TES 7  
|||  
Db 5 TES 7

#### RESULT 12

A36527

juvenile-hormone esterase (EC 3.1.1.59) - tobacco hornworm (fragment)

C;Species: Manduca sexta (tobacco hornworm)

C;Date: 12-Apr-1991 #sequence\_revision 12-Apr-1991 #text\_change 03-Feb-1994

C;Accession: A36527

R;Venkatesh, K.; Abdel-Aal, Y.A.I.; Armstrong, F.B.; Roe, R.M.

J. Biol. Chem. 265, 21727-21732, 1990

A;Title: Characterization of affinity-purified juvenile hormone esterase from the plasma of the tobacco hornworm, Manduca sexta.

A;Reference number: A36527; MUID:91072375; PMID:2254326

A;Accession: A36527

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-15 <VEN>

C;Keywords: carboxylic ester hydrolase

Query Match 37.5%; Score 3; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TES 7  
|||  
Db 12 TES 14

#### RESULT 13

A49480

major immunophilin hsp56 - chicken (fragment)

C;Species: Gallus gallus (chicken)

C;Date: 06-Oct-1994 #sequence\_revision 18-Nov-1994 #text\_change 28-Apr-1995

C;Accession: A49480

R;Yem, A.W.; Reardon, I.M.; Leone, J.W.; Heinrikson, R.L.; Deibel Jr., M.R.

Biochemistry 32, 12571-12576, 1993

A;Title: An active FK506-binding domain of 17,000 daltons is isolated following limited proteolysis of chicken thymus hsp56.

A;Reference number: A49480; MUID:94072550; PMID:7504525

A;Accession: A49480

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-15 <YEM>

A;Experimental source: thymus

A;Note: sequence extracted from NCBI backbone (NCBIP:142438)

Query Match 37.5%; Score 3; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy            5 TES 7  
             |||  
Db            7 TES 9

RESULT 14

B45115

peptidylprolyl isomerase (EC 5.2.1.8) FKBP51 - human (fragment)

N;Alternate names: FK506-binding protein FKBP51; peptidylprolyl cis-trans isomerase FKBP51

C;Species: Homo sapiens (man)

C;Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 16-Feb-1997

C;Accession: B45115

R;Wiederrecht, G.; Hung, S.; Chan, H.K.; Marcy, A.; Martin, M.; Calaycay, J.; Boulton, D.; Sigal, N.; Kincaid, R.L.; Siekierka, J.J.

J. Biol. Chem. 267, 21753-21760, 1992

A;Title: Characterization of high molecular weight FK-506 binding activities reveals a novel FK-506-binding protein as well as a protein complex.

A;Reference number: A45115; MUID:93016131; PMID:1383226

A;Accession: B45115

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-15 <WIE>

A;Experimental source: JURKAT cells

A;Note: sequence extracted from NCBI backbone (NCBIP:116748)

C;Keywords: cis-trans-isomerase; cyclosporin A binding

Query Match                    37.5%;   Score 3;   DB 2;   Length 15;  
Best Local Similarity       100.0%;   Pred. No. 1.7e+03;  
Matches       3;   Conservative       0;   Mismatches       0;   Indels       0;   Gaps       0;

Qy            5 TES 7  
             |||  
Db            2 TES 4

RESULT 15

S05671

hirudin Ia - medicinal leech (fragment)

N;Alternate names: thrombin inhibitor

C;Species: Hirudo medicinalis (medicinal leech)

C;Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 17-Jul-1998

C;Accession: S05671

R;Scharf, M.; Engels, J.; Tripier, D.

FEBS Lett. 255, 105-110, 1989

A;Title: Primary structures of new 'iso-hirudins'.

A;Reference number: S05671; MUID:90005945; PMID:2792365

A;Accession: S05671

A;Molecule type: protein

A;Residues: 1-17 <SCH>

C;Superfamily: thrombin inhibitor

C;Keywords: anticoagulant; serine proteinase inhibitor; sulfoprotein

F;6-14/Disulfide bonds: #status predicted

Query Match                    37.5%;   Score 3;   DB 2;   Length 17;  
Best Local Similarity       100.0%;   Pred. No. 1.9e+03;  
Matches       3;   Conservative       0;   Mismatches       0;   Indels       0;   Gaps       0;



Qy	5	TES	7
Db	7	TES	9

Search completed: November 13, 2003, 10:39:56  
Job time : 7.83333 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:55:06 ; Search time 4.33333 Seconds  
(without alignments)  
86.819 Million cell updates/sec

```
Title:      US-09-228-866-9
Perfect score: 8
Sequence:   1 CTRITESC 8
```

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1246

```
Minimum DB seq length: 7
Maximum DB seq length: 21
```

Post-processing: Listing first 45 summaries

Database : SwissProt 41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	%				Query	DB	ID	Description
	No.	Score	Match	Length				
1	3	37.5	7	1	IGAO_DACDE	P06294	dactylum d	
2	3	37.5	10	1	MALE_KLEPN	Q05564	klebsiella	
3	3	37.5	15	1	ESTB_SCHGA	P81011	schizaphis	
4	3	37.5	15	1	ESTJ_MANSE	P19985	manduca sex	
5	3	37.5	16	1	FIBA_MANLE	P14455	mandrillus	
6	3	37.5	16	1	H5_COTJA	P18638	coturnix co	
7	3	37.5	17	1	ITHB_HIRME	P28502	hirudo medi	
8	3	37.5	21	1	ATPB_PHYPA	P80658	physcomitre	
9	2	25.0	8	1	UPA1 HUMAN	P30087	homo sapien	

10	2	25.0	9	1	DNF1_LOCFI	P16339	locusta mig
11	2	25.0	9	1	FIBB_MACFU	P19345	macaca fusc
12	2	25.0	9	1	IPYR_RHOVI	P82992	rhodopseudo
13	2	25.0	9	1	OXYT_BUFRE	P42995	bufo regula
14	2	25.0	9	1	RS10_SERMA	O68936	serratia ma
15	2	25.0	9	1	UHA2_HUMAN	P40929	homo sapien
16	2	25.0	9	1	UN19_CLOPA	P81355	clostridium
17	2	25.0	10	1	COXQ_SHEEP	P80337	ovis aries
18	2	25.0	10	1	ESL_LACCA	P81758	lactobacill
19	2	25.0	10	1	UHA3_HUMAN	P40930	homo sapien
20	2	25.0	11	1	ANGT_CRIGE	P09037	crinia geor
21	2	25.0	11	1	ASL2_BACSE	P83147	bacteroides
22	2	25.0	11	1	RR2_CONAM	P42341	conopholis
23	2	25.0	11	1	TIN1_HOPTI	P82651	hoplobatrac
24	2	25.0	12	1	CALM_TETTH	Q05055	tetrahymena
25	2	25.0	12	1	CXST_CONTE	P58846	conus texti
26	2	25.0	12	1	PA21_MICFM	P25072	micrurus fu
27	2	25.0	12	1	PSP3_PHYPA	P80662	physcomitre
28	2	25.0	12	1	UN39_CLOPA	P81359	clostridium
29	2	25.0	12	1	UR2B_CATCO	P04559	catostomus
30	2	25.0	12	1	UR2B_CYPCA	P04561	cyprinus ca
31	2	25.0	13	1	AU11_LITRA	P82386	litoria ran
32	2	25.0	13	1	AU12_LITRA	P82387	litoria ran
33	2	25.0	13	1	BP37_LEUMA	P81754	leucophaea
34	2	25.0	13	1	CXA2_CONGE	P01520	conus geogr
35	2	25.0	13	1	FIBA_CAVPO	P14445	cavia porce
36	2	25.0	13	1	IDHP_RAT	P56574	rattus norv
37	2	25.0	13	1	NP1_LYMST	P80178	lymnaea sta
38	2	25.0	13	1	NP2_LYMST	P80179	lymnaea sta
39	2	25.0	13	1	NP3_LYMST	P80180	lymnaea sta
40	2	25.0	13	1	NP4_LYMST	P80181	lymnaea sta
41	2	25.0	13	1	NP5_LYMST	P80182	lymnaea sta
42	2	25.0	13	1	RPOC_MYCGA	P47716	mycoplasma
43	2	25.0	13	1	TEML_RANTE	P57104	rana tempor
44	2	25.0	13	1	UHA3_CANFA	P56535	canis famil
45	2	25.0	13	1	UN02_PINPS	P81667	pinus pinas

# ALIGNMENTS

## RESULT 1

IGAO\_DACDE

ID IGAO\_DACDE STANDARD; PRT; 7 AA.

AC P06294;

DT 01-JAN-1988 (Rel. 06, Created)

DT 01-JAN-1988 (Rel. 06, Last sequence update)

DT 01-OCT-1994 (Rel. 30, Last annotation update)

DE Galactose oxidase inhibitor.

OS Dactylium dendroides (Cladobotryum dendroides).

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;

OC Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypomyces.

OX NCBI\_TaxID=5132;

RN [1]

RP SEQUENCE.

RA Avigad G., Markus Z.;

RT "Identification of a peptide inhibitor of galactose oxidase from

RT Dactylium dendroides.";  
 RL Fed. Proc. 31:447-447(1972).  
 CC -!- FUNCTION: BINDS ONE COPPER ION PER MOLECULE BUT DOES NOT BIND THE  
 CC GALACTOSE OXIDASE APOENZYME. IT MAY INACTIVATE THE ENZYME BY  
 CC BINDING TO ITS PROSTHETIC COPPER GROUP.  
 DR PIR; A01341; XEYDGD.  
 KW Copper; Metalloenzyme inhibitor.  
 SQ SEQUENCE 7 AA; 706 MW; 75BB01A456D87DB0 CRC64;

Query Match 37.5%; Score 3; DB 1; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TES 7  
 |||  
 Db 5 TES 7

## RESULT 2

### MALE\_KLEPN

ID MALE\_KLEPN STANDARD; PRT; 10 AA.  
 AC Q05564;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Maltose-binding periplasmic protein (Maltodextrin-binding protein)  
 DE (MMBP) (Fragment).  
 GN MALE.  
 OS Klebsiella pneumoniae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Klebsiella.  
 OX NCBI\_TaxID=573;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1033-5P14 / KAY2026;  
 RX MEDLINE=93211295; PubMed=8459773;  
 RA Bachellier S., Perrin D., Hofnung M., Gilson E.;  
 RT "Bacterial interspersed mosaic elements (BIMEs) are present in the  
 RT genome of Klebsiella.";  
 RL Mol. Microbiol. 7:537-544(1993).  
 CC -!- FUNCTION: MALE IS INVOLVED IN THE HIGH-AFFINITY MALTOSE MEMBRANE  
 CC TRANSPORT SYSTEM. INITIAL RECEPTOR FOR THE ACTIVE TRANSPORT OF AND  
 CC CHEMOTAXIS TOWARD MALTOOLIGOSACCHARIDES.  
 CC -!- SUBCELLULAR LOCATION: Periplasmic.  
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING  
 CC PROTEIN FAMILY 1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X68329; CAA48406.1; -.  
 DR InterPro; IPR006061; SBP\_dom1.

DR PROSITE; PS01037; SBP\_BACTERIAL\_1; PARTIAL.  
KW Transport; Sugar transport; Periplasmic.  
FT NON\_TER 1 1  
SQ SEQUENCE 10 AA; 1159 MW; 8FD8DC4415A6DDDA CRC64;

Query Match 37.5%; Score 3; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RIT 5  
|||  
Db 7 RIT 9

#### RESULT 3

##### ESTB\_SCHGA

ID ESTB\_SCHGA STANDARD; PRT; 15 AA.  
AC P81011;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Esterase 56 kDa subunit (EC 3.1.1.1) (Carboxylic-ester hydrolase)  
DE (Fragment).  
OS Schizaphis graminum (Aphid).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;  
OC Aphidoidea; Aphididae; Aphidini; Schizaphis.  
OX NCBI\_TaxID=13262;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=97468499; PubMed=9327586;  
RA Siegfried B.D., Ono M., Swanson J.J.;  
RT "Purification and characterization of a carboxylesterase associated  
RT with organophosphate resistance in the greenbug, Schizaphis graminum  
RT (Homoptera: Aphididae).";  
RL Arch. Insect Biochem. Physiol. 36:229-240(1997).  
CC -!- CATALYTIC ACTIVITY: A carboxylic ester + H(2)O = an alcohol + a  
CC carboxylic anion.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.  
DR InterPro; IPR002018; CarbesteraseB.  
DR PROSITE; PS00122; CARBOXYLESTERASE\_B\_1; PARTIAL.  
DR PROSITE; PS00941; CARBOXYLESTERASE\_B\_2; PARTIAL.  
KW Hydrolase; Serine esterase.  
FT NON\_TER 15 15  
SQ SEQUENCE 15 AA; 1609 MW; 1208B2BCCC969482 CRC64;

Query Match 37.5%; Score 3; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 5.1e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RIT 5  
|||  
Db 5 RIT 7

#### RESULT 4

## ESTJ\_MANSE

ID ESTJ\_MANSE STANDARD; PRT; 15 AA.  
 AC P19985;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE Juvenile hormone esterase (EC 3.1.1.59) (JH esterase) (Fragment).  
 OS *Manduca sexta* (Tobacco hawkmoth) (Tobacco hornworm).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingioidea;  
 OC Sphingidae; Sphinginae; *Manduca*.  
 OX NCBI\_TaxID=7130;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Larval plasma;  
 RX MEDLINE=91072375; PubMed=2254326;  
 RA Venkatesh K., Abdel-Aal Y.A.I., Armstrong F.B., Roe R.M.;  
 RT "Characterization of affinity-purified juvenile hormone esterase from  
 RT the plasma of the tobacco hornworm, *Manduca sexta*."  
 RL J. Biol. Chem. 265:21727-21732(1990).  
 CC -!- FUNCTION: JH ESTERASE PLAYS A CRUCIAL ROLE IN THE DECREASE OF  
 CC JH ACTIVITY IN LEPIDOPTERAN INSECTS, BY HYDROLYZING THE METHYL  
 CC ESTER OF JH. IT IS ALSO INVOLVED IN THE TRANSPORT OF JH.  
 CC -!- CATALYTIC ACTIVITY: Methyl (2E,6E)-(10R,11S)-10,11-epoxy-3,7,11-  
 CC trimethyltrideca-2,6-dienoate + H(2)O = (2E,6E)-(10R,11S)-10,11-  
 CC epoxy-3,7,11-trimethyltrideca-2,6-dienoate + methanol.  
 CC -!- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.  
 DR PIR; A36527; A36527.  
 DR InterPro; IPR002018; CarbesteraseB.  
 DR PROSITE; PS00122; CARBOXYLESTERASE\_B\_1; PARTIAL.  
 DR PROSITE; PS00941; CARBOXYLESTERASE\_B\_2; PARTIAL.  
 KW Hydrolase; Serine esterase.  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1659 MW; D321EA432E58B848 CRC64;

Query Match 37.5%; Score 3; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TES 7  
 |||  
 Db 12 TES 14

## RESULT 5

## FIBA\_MANLE

ID FIBA\_MANLE STANDARD; PRT; 16 AA.  
 AC P14455;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).  
 GN FGA.  
 OS *Mandrillus leucophaeus* (Drill) (*Papio leucophaeus*).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; *Mandrillus*.

OX NCBI\_TaxID=9568;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=69115139; PubMed=4974768;  
 RA Doolittle R.F., Glasgow C., Mross G.A.;  
 RT "Characterization of fibrinopeptides A and B from a drill (Mandrillus  
 RT leucophaeus).";  
 RL Biochim. Biophys. Acta 175:217-219(1969).  
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.  
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 KW Blood coagulation; Plasma.  
 FT PEPTIDE 1 16 FIBRINOPEPTIDE A.  
 FT NON\_TER 16 16  
 SQ SEQUENCE 16 AA; 1567 MW; 08E8CBB77BA051A4 CRC64;

Query Match 37.5%; Score 3; DB 1; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITE 6  
 |||  
 Db 9 ITE 11

#### RESULT 6

##### H5\_COTJA

ID H5\_COTJA STANDARD; PRT; 16 AA.  
 AC P18638;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Histone H5 (Fragment).  
 OS Coturnix coturnix japonica (Japanese quail).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Coturnix.  
 OX NCBI\_TaxID=93934;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=76277939; PubMed=962913;  
 RA Seligy V., Roy C., Dove M., Yaguchi M.;  
 RT "Species variability of N-terminal sequence of avian erythrocyte-  
 RT specific histone H5.";  
 RL Biochem. Biophys. Res. Commun. 71:196-202(1976).  
 CC -!- FUNCTION: HISTONE H5 PERFORMS THE SAME FUNCTION AS H1, BEING  
 CC NECESSARY FOR THE CONDENSATION OF NUCLEOSOME CHAINS INTO HIGHER  
 CC ORDER STRUCTURES, AND REPLACES HISTONE H1 IN CERTAIN CELLS.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- TISSUE SPECIFICITY: ERYTHROID CELLS.  
 CC -!- SIMILARITY: BELONGS TO THE HISTONE H1/H5 FAMILY.

KW Chromosomal protein; Nuclear protein; DNA-binding; DNA condensation.  
FT NON\_TER 16 16  
SQ SEQUENCE 16 AA; 1665 MW; DB528219B3074D3C CRC64;

Query Match 37.5%; Score 3; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 5.4e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TES 7  
|||  
Db 1 TES 3

#### RESULT 7

##### ITHB\_HIRME

ID ITHB\_HIRME STANDARD; PRT; 17 AA.  
AC P28502;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hirudin IA (Fragment).  
OS Hirudo medicinalis (Medicinal leech).  
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;  
OC Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.  
OX NCBI\_TaxID=6421;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=90005945; PubMed=2792365;  
RA Scharf M., Engels J., Tripiet D.;  
RT "Primary structures of new 'iso-hirudins'.";  
RL FEBS Lett. 255:105-110(1989).  
CC -!- FUNCTION: Hirudin is a potent thrombin-specific protease  
CC inhibitor. It forms a stable non-covalent complex with alpha-  
CC thrombin, thereby abolishing its ability to cleave fibrinogen.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: BELONGS TO THE HIRUDIN FAMILY.  
DR PIR; S05671; S05671.  
DR HSSP; P01050; 1HIC.  
DR InterPro; IPR000429; Hirudin.  
DR Pfam; PF00713; Hirudin; 1.  
KW Serine protease inhibitor; Sulfation; Multigene family.  
FT NON\_TER 17 17  
SQ SEQUENCE 17 AA; 1877 MW; 8904C6786C301CE7 CRC64;

Query Match 37.5%; Score 3; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 5.7e+02;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TES 7  
|||  
Db 7 TES 9

#### RESULT 8

##### ATPB\_PHYPA

ID ATPB\_PHYPA STANDARD; PRT; 21 AA.  
AC P80658;

DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ATP synthase beta chain (EC 3.6.3.14) (Fragment).  
 GN ATPB.  
 OS Physcomitrella patens (Moss).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;  
 OC Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.  
 OX NCBI\_TaxID=3218;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Protonema;  
 RX MEDLINE=97275459; PubMed=9129336;  
 RA Kasten B., Buck F., Nuske J., Reski R.;  
 RT "Cytokinin affects nuclear- and plastome-encoded energy-converting  
 RT plastid enzymes.";  
 RL Planta 201:261-272(1997).  
 CC -!- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON  
 CC GRADIENT ACROSS THE MEMBRANE. THE BETA CHAIN IS THE CATALYTIC  
 CC SUBUNIT.  
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +  
 CC H(+) (Out).  
 CC -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC  
 CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE  
 CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)  
 CC HAS THREE MAIN SUBUNITS: A, B AND C.  
 CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.  
 CC -!- SIMILARITY: Belongs to the ATPase alpha/beta chains family.  
 DR InterPro; IPR000194; ATPase\_a/bcentre.  
 DR PROSITE; PS00152; ATPASE\_ALPHA\_BETA; PARTIAL.  
 KW ATP synthesis; Chloroplast; Thylakoid; Membrane; CF(1);  
 KW Hydrolase; ATP-binding; Hydrogen ion transport.  
 FT NON\_TER 21 21  
 SQ SEQUENCE 21 AA; 2298 MW; 9558E4F5AC89D81A CRC64;

Query Match 37.5%; Score 3; DB 1; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RIT 5  
 |||  
 Db 7 RIT 9

#### RESULT 9

##### UPA1\_HUMAN

ID UPA1\_HUMAN STANDARD; PRT; 8 AA.  
 AC P30087;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Unknown protein from 2D-page of plasma (Spot 2) (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;



RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Plasma;  
 RX MEDLINE=93092937; PubMed=1459097;  
 RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,  
 RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
 RA Hochstrasser D.F.;  
 RT "Plasma protein map: an update by microsequencing";  
 RL Electrophoresis 13:707-714(1992).  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN  
 CC PROTEIN IS: 4.9, ITS MW IS: 65 kDa.  
 DR SWISS-2DPAGE; P30087; HUMAN.  
 FT NON\_TER 1 1  
 FT UNSURE 8 8  
 FT NON\_TER 8 8  
 SQ SEQUENCE 8 AA; 944 MW; C01772C455BB06DA CRC64;

Query Match 25.0%; Score 2; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ES 7  
 ||  
 Db 3 ES 4

# RESULT 10

## DNF1\_LOCF1

ID DNF1\_LOCF1 STANDARD; PRT; 9 AA.  
 AC P16339;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Locupressin (Diuretic neuropeptide F1/F2).  
 OS Locusta migratoria (Migratory locust).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;  
 OC Acridoidea; Acrididae; Oedipodinae; Locusta.  
 OX NCBI\_TaxID=7004;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Suboesophageal ganglion, and Thoracic ganglion;  
 RX MEDLINE=88077077; PubMed=3689410;  
 RA Proux J.P., Miller C.A., Li J.P., Carney R.L., Girardie A.,  
 RA Delaage M., Schooley D.A.;  
 RT "Identification of an arginine vasopressin-like diuretic hormone from  
 RT Locusta migratoria";  
 RL Biochem. Biophys. Res. Commun. 149:180-186(1987).  
 CC -!- FUNCTION: DIURETIC HORMONE.  
 CC -!- SUBUNIT: F2 IS AN ANTIPARALLEL DISULFIDE LINKED DIMER OF F1.  
 CC -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.  
 DR PIR; A29477; A29477.  
 DR InterPro; IPR000981; Neurhyp\_horm.  
 DR Pfam; PF00220; hormone4; 1.  
 DR PROSITE; PS00264; NEUROHYPOPHYS\_HORM; 1.  
 KW Hormone; Neuropeptide; Amidation.  
 FT DISULFID 1 6 IN F1.

FT DISULFID 1 1 INTERCHAIN (WITH C-6') (IN F2).  
 FT DISULFID 6 6 INTERCHAIN (WITH C-1') (IN F2).  
 FT MOD\_RES 9 9 AMIDATION.  
 SQ SEQUENCE 9 AA; 976 MW; 56EB176EB451A057 CRC64;

Query Match 25.0%; Score 2; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 IT 5  
 ||  
 Db 3 IT 4

# RESULT 11

## FIBB\_MACFU

ID FIBB\_MACFU STANDARD; PRT; 9 AA.  
 AC P19345;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).  
 GN FGB.  
 OS Macaca fuscata fuscata (Japanese macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9543;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=85289140; PubMed=3928610;  
 RA Nakamura S., Takenaka O., Takahashi K.;  
 RT "Fibrinopeptides A and B of Japanese monkey (Macaca fuscata) and  
 RT patas monkey (Erythrocebus patas): their amino acid sequences,  
 RT restricted mutations, and a molecular phylogeny for macaques,  
 RT guenons, and baboons.";  
 RL J. Biochem. 97:1487-1492(1985).  
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.  
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 DR PIR; C24180; C24180.  
 DR InterPro; IPR002181; Fibrinogen\_C.  
 DR PROSITE; PS00514; FIBRIN\_AG\_C\_DOMAIN; PARTIAL.  
 KW Blood coagulation; Plasma.  
 FT PEPTIDE 1 9 FIBRINOPEPTIDE B.  
 FT NON\_TER 9 9  
 SQ SEQUENCE 9 AA; 1038 MW; 69FE65B9C735BB1B CRC64;

Query Match 25.0%; Score 2; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy            6 ES 7  
               ||  
 Db            3 ES 4

RESULT 12

IPYR\_RHOVI

ID IPYR\_RHOVI            STANDARD;            PRT;            9 AA.  
 AC P82992;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-  
 DE hydrolase) (PPase) (Fragment).  
 GN PPA.  
 OS Rhodopseudomonas viridis.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Hyphomicrobiaceae; Blastochloris.  
 OX NCBI\_TaxID=1079;  
 RN [1]  
 RP SEQUENCE, FUNCTION, AND SUBCELLULAR LOCATION.  
 RA Gomez R., Losada M., Serrano A.;  
 RL Submitted (JUN-2001) to the SWISS-PROT data bank.  
 CC -!- CATALYTIC ACTIVITY: Diphosphate + H(2)O = 2 phosphate.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED MW OF THIS PROTEIN IS:  
 CC 23 kDa.  
 CC -!- SIMILARITY: Belongs to the PPase family.  
 DR HAMAP; MF\_00209; -; 1.  
 DR InterPro; IPR001596; Pyrophosphatase.  
 DR PROSITE; PS00387; PPASE; PARTIAL.  
 KW Hydrolase.  
 FT NON\_TER            9            9  
 SQ SEQUENCE    9 AA; 1014 MW; 68F3EAA05DDAA044 CRC64;

Query Match            25.0%; Score 2; DB 1; Length 9;  
 Best Local Similarity    100.0%; Pred. No. 1.3e+05;  
 Matches    2; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

Qy            3 RI 4  
               ||  
 Db            2 RI 3

RESULT 13

OXYT\_BUFRE

ID OXYT\_BUFRE            STANDARD;            PRT;            9 AA.  
 AC P42995;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Seritocin ([Ser5,Ile8]-oxytocin).  
 OS Bufo regularis (Leopard toad).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Bufonidae;  
 OC Bufo.

OX NCBI\_TaxID=8390;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Pituitary neurointermediate lobe;  
 RX MEDLINE=96059313; PubMed=7591488;  
 RA Chauvet J., Michel G., Ouedraogo Y., Chou J., Chait B.T., Acher R.;  
 RT "A new neurohypophysial peptide, seritocin ([Ser5,Ile8]-oxytocin),  
 RT identified in a dryness-resistant African toad, Bufo regularis."  
 RL Int. J. Pept. Protein Res. 45:482-487(1995).  
 CC -!- FUNCTION: Devoid of oxytocic activity.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.  
 DR InterPro; IPR000981; Neurhyp\_horm.  
 DR Pfam; PF00220; hormone4; 1.  
 DR PROSITE; PS00264; NEUROHYPOPHYS\_HORM; FALSE\_NEG.  
 KW Hormone; Amidation.  
 FT DISULFID 1 6  
 FT MOD\_RES 9 9 AMIDATION.  
 SQ SEQUENCE 9 AA; 983 MW; 17FF476EA5A6D04B CRC64;

Query Match 25.0%; Score 2; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 SC 8  
 ||  
 Db 5 SC 6

#### RESULT 14

##### RS10\_SERMA

ID RS10\_SERMA STANDARD; PRT; 9 AA.  
 AC O68936;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 30S ribosomal protein S10 (Fragment).  
 GN RPSJ.  
 OS Serratia marcescens.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Serratia.  
 OX NCBI\_TaxID=615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Noorani S.M., Lindahl L., Zengel J.M.;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: Involved in the binding of tRNA to the ribosomes (By  
 CC similarity).  
 CC -!- SIMILARITY: BELONGS TO THE S10P FAMILY OF RIBOSOMAL PROTEINS.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC -----
DR   EMBL; AF058451; AAC14294.1; -.
DR   HAMAP; MF_00508; -; 1.
DR   InterPro; IPR001848; Ribosomal_S10.
DR   PROSITE; PS00361; RIBOSOMAL_S10; PARTIAL.
KW   Ribosomal protein.
FT   NON_TER      9      9
SQ   SEQUENCE     9 AA; 1214 MW; DE3944004416D456 CRC64;

```

```

Query Match      25.0%; Score 2; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches      2; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

QY      3 RI 4
      ||
Db      5 RI 6

```

# RESULT 15

## UHA2\_HUMAN

```

ID   UHA2_HUMAN      STANDARD;          PRT;          9 AA.
AC   P40929;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Unknown protein from 2D-page of heart (Spot 5603) (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Heart;
RX   MEDLINE=95203287; PubMed=7895732;
RA   Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT   "The human myocardial two-dimensional gel protein database: update
RT   1994.";
RL   Electrophoresis 15:1459-1465(1994).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC   PROTEIN IS: 6.0, ITS MW IS: 55.3 kDa.
FT   NON_TER      9      9
SQ   SEQUENCE     9 AA; 1104 MW; 8874B1BB5B01B2CA CRC64;

```

```

Query Match      25.0%; Score 2; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches      2; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

QY      5 TE 6
      ||
Db      4 TE 5

```

Search completed: November 13, 2003, 10:34:00  
Job time : 4.33333 secs

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:58:36 ; Search time 19.6667 Seconds  
(without alignments)  
104.971 Million cell updates/sec

Title: US-09-228-866-9  
Perfect score: 8  
Sequence: 1 CTRITESC 8

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 7516

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4	50.0	20	6	Q9BEC6	Q9bec6 sus scrofa
2	3	37.5	9	15	Q8AEW8	Q8aew8 human immun
3	3	37.5	10	6	Q9XS84	Q9xs84 equus cabal

4	3	37.5	11	2	Q93MI7	Q93mi7 escherichia
5	3	37.5	11	4	Q9H3V7	Q9h3v7 homo sapien
6	3	37.5	12	2	Q47251	Q47251 escherichia
7	3	37.5	12	4	Q13695	Q13695 homo sapien
8	3	37.5	12	11	Q9WUX1	Q9wux1 rattus norv
9	3	37.5	13	2	Q52543	Q52543 pseudomonas
10	3	37.5	13	2	Q53693	Q53693 streptomyce
11	3	37.5	13	4	Q9NR93	Q9nr93 homo sapien
12	3	37.5	13	10	Q8SAT3	Q8sat3 flaveria ro
13	3	37.5	13	10	Q9S922	Q9s922 brassica ol
14	3	37.5	13	10	Q8SAT4	Q8sat4 flaveria an
15	3	37.5	15	13	Q9PRM3	Q9prm3 gallus gall
16	3	37.5	16	2	Q9EVA6	Q9eva6 aphanizomen
17	3	37.5	16	2	Q9R4J4	Q9r4j4 pseudomonas
18	3	37.5	16	2	Q8KLP7	Q8klp7 streptomyce
19	3	37.5	16	2	Q9EVA8	Q9eva8 aphanizomen
20	3	37.5	16	2	Q9EVB0	Q9evb0 aphanizomen
21	3	37.5	16	2	Q9EVA4	Q9eva4 aphanizomen
22	3	37.5	16	6	Q9TRJ5	Q9trj5 bos taurus
23	3	37.5	17	8	Q36741	Q36741 homo sapien
24	3	37.5	18	2	Q9R5F9	Q9r5f9 alcaligenes
25	3	37.5	18	8	Q95AV2	Q95av2 fucus serra
26	3	37.5	18	8	Q9ZY75	Q9zy75 cenocoelius
27	3	37.5	18	8	Q95AV3	Q95av3 fucus evane
28	3	37.5	18	10	Q9S8I7	Q9s8i7 oryza sativ
29	3	37.5	18	10	Q9S8I8	Q9s8i8 oryza sativ
30	3	37.5	18	15	Q9PXF1	Q9pxf1 human immun
31	3	37.5	19	8	Q9GI99	Q9gi99 sargassum t
32	3	37.5	19	10	Q9S923	Q9s923 brassica ol
33	3	37.5	19	12	Q84864	Q84864 unidentifie
34	3	37.5	19	15	Q905K3	Q905k3 human immun
35	3	37.5	20	2	P83072	P83072 bacillus ce
36	3	37.5	20	2	P83162	P83162 anabaena sp
37	3	37.5	20	5	P82109	P82109 mythimna un
38	3	37.5	20	8	Q9BC69	Q9bc69 sargassum s
39	3	37.5	20	10	Q9S8M3	Q9s8m3 solanum tub
40	3	37.5	20	11	Q9QV46	Q9qv46 rattus sp.
41	3	37.5	20	11	Q924T0	Q924t0 rattus norv
42	3	37.5	21	2	Q9R4U0	Q9r4u0 streptomyce
43	3	37.5	21	4	Q96PP2	Q96pp2 homo sapien
44	3	37.5	21	8	Q9GIA7	Q9gia7 sargassum p
45	3	37.5	21	10	Q42734	Q42734 flaveria bi

# ALIGNMENTS

## RESULT 1

Q9BEC6

ID Q9BEC6 PRELIMINARY; PRT; 20 AA.

AC Q9BEC6;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE Peroxisomal membrane protein 1 (Fragment).

GN ABCD3.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Pietrain;  
 RA Stratil A., Kubickova S., Peelman L.J., Reiner G., Musilova P.,  
 RA Van Poucke M., Rubes J., Geldermann H.;  
 RT "FISH, RH and linkage assignment of the porcine ABCD3 (PXMP1) gene to  
 RT the distal end of chromosome 4q.";  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AJ309827; CAC32854.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 20 20  
 SQ SEQUENCE 20 AA; 2215 MW; A97F37E52C8D5DF8 CRC64;

Query Match 50.0%; Score 4; DB 6; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RITE 6  
 ||||  
 Db 6 RITE 9

## RESULT 2

Q8AEW8

ID Q8AEW8 PRELIMINARY; PRT; 9 AA.  
 AC Q8AEW8;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE REV protein (Anti-repression transactivator protein) (Fragment).  
 GN REV.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HAART patient 26;  
 RA Saurya S.;  
 RT "Characterization of HIV-1 genes from AIDS patients on combination  
 RT therapy with discordance between viral load and CD4+ T cell counts.";  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AJ496724; CAD43154.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 9 AA; 1020 MW; 9D2A4411ADC1AB05 CRC64;

Query Match 37.5%; Score 3; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 8.3e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TRI 4  
 |||  
 Db 6 TRI 8



# RESULT 3

Q9XS84

ID Q9XS84 PRELIMINARY; PRT; 10 AA.  
AC Q9XS84;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Leptin (Fragment).  
GN LEP.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99160468; PubMed=10051323;  
RA Caetano A.R., Pomp D., Murray J.D., Bowling A.T.;  
RT "Comparative mapping of 18 equine type I genes assigned by somatic  
cell hybrid analysis."  
RL Mamm. Genome 10:271-276(1999).  
DR EMBL; AF097582; AAD25985.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 10 10  
SQ SEQUENCE 10 AA; 1155 MW; 6BE77BB05AA44044 CRC64;

Query Match 37.5%; Score 3; DB 6; Length 10;  
Best Local Similarity 100.0%; Pred. No. 2.9e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TRI 4  
|||  
Db 2 TRI 4

# RESULT 4

Q93MI7

ID Q93MI7 PRELIMINARY; PRT; 11 AA.  
AC Q93MI7;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Adhesin (Fragment).  
GN IHA.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CFT073;  
RA Stell A.L.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF401752; AAK94916.1; -.  
FT NON\_TER 11 11  
SQ SEQUENCE 11 AA; 1203 MW; 8E2817ECBDD731B1 CRC64;

Query Match 37.5%; Score 3; DB 2; Length 11;

Best Local Similarity 100.0%; Pred. No. 3.1e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RIT 5

|||

Db 2 RIT 4

#### RESULT 5

Q9H3V7

ID Q9H3V7 PRELIMINARY; PRT; 11 AA.  
AC Q9H3V7;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE Aquaporin-4 (Fragment).  
GN AQP4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97008105; PubMed=8855281;  
RA Lu M., Lee M.D., Smith B.L., Jung J.S., Agre P., Verdijk M.A.J.,  
RA Merckx G., Rijss J.P.L., Deen P.M.T.;  
RT "The human AQP4 gene: definition of the locus encoding two water  
RT channel polypeptides in brain."  
RL Proc. Natl. Acad. Sci. U.S.A. 93:10908-10912(1996).  
DR EMBL; U63613; AAG17964.1; -.  
KW Porin.  
FT NON\_TER 1 1  
FT NON\_TER 11 11  
SQ SEQUENCE 11 AA; 1233 MW; 379D9CA311AEB737 CRC64;

Query Match 37.5%; Score 3; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 3.1e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTR 3

|||

Db 6 CTR 8

#### RESULT 6

Q47251

ID Q47251 PRELIMINARY; PRT; 12 AA.  
AC Q47251;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE HEMC protein (Fragment).  
GN HEMC.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=88096587; PubMed=3320969;  
 RA Jordan P.M., Mgbeje B.I.A., Alwan A.F., Thomas S.D.;  
 RT "Nucleotide sequence of hemD, the second gene in the hem operon of  
 RT Escherichia coli K-12."  
 RL Nucleic Acids Res. 15:10583-10583(1987).  
 DR EMBL; Y00883; CAA68775.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1375 MW; 5D2DE8339BA045B3 CRC64;

Query Match 37.5%; Score 3; DB 2; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITE 6  
 |||  
 Db 7 ITE 9

#### RESULT 7

Q13695

ID Q13695 PRELIMINARY; PRT; 12 AA.  
 AC Q13695;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Acetylcholine receptor (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Thymus;  
 RX MEDLINE=94071933; PubMed=8250918;  
 RA Mihovilovic M., Mai Y., Herbstreith M., Rubboli F., Tarroni P.,  
 RA Clementi F., Roses A.D.;  
 RT "Splicing of an anti-sense Alu sequence generates a coding sequence  
 RT variant for the alpha-3 subunit of a neuronal acetylcholine  
 RT receptor."  
 RL Biochem. Biophys. Res. Commun. 197:137-144(1993).  
 DR EMBL; L18973; AAA86792.1; -.  
 KW Receptor.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 12 AA; 1282 MW; 8B0B423A6FC1B2D5 CRC64;

Query Match 37.5%; Score 3; DB 4; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ESC 8  
 |||  
 Db 2 ESC 4

# RESULT 8

Q9WUX1

ID Q9WUX1 PRELIMINARY; PRT; 12 AA.  
AC Q9WUX1;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)  
DE APOBEC-1 protein (Fragment).  
GN APOBEC-1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98335789; PubMed=9672068;  
RA Greeve J., Axelos D., Welker S., Schipper M., Greten H.;  
RT "Distinct promoters induce APOBEC-1 expression in rat liver and  
RT intestine.";  
RL Arterioscler. Thromb. Vasc. Biol. 18:1079-1092(1998).  
DR EMBL; AJ006695; CAB44439.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 12 12  
SQ SEQUENCE 12 AA; 1357 MW; 70FB1679699325BB CRC64;

Query Match 37.5%; Score 3; DB 11; Length 12;  
Best Local Similarity 100.0%; Pred. No. 3.3e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TES 7  
| | |  
Db 4 TES 6

# RESULT 9

Q52543

ID Q52543 PRELIMINARY; PRT; 13 AA.  
AC Q52543;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE Pilus expression protein (Fragment).  
GN PILQ.  
OS Pseudomonas syringae.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=317;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HB10Y;  
RX MEDLINE=96134973; PubMed=8550460;  
RA Roine E., Nunn D.N., Paulin L., Romantschuk M.;  
RT "Characterization of genes required for pilus expression in  
RT Pseudomonas syringae pathovar phaseolicola.";  
RL J. Bacteriol. 178:410-417(1996).  
DR EMBL; L28837; AAA93090.1; -.

FT NON\_TER 13 13  
SQ SEQUENCE 13 AA; 1406 MW; 9F915F7893495045 CRC64;

Query Match 37.5%; Score 3; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TRI 4  
|||  
Db 2 TRI 4

RESULT 10

Q53693

ID Q53693 PRELIMINARY; PRT; 13 AA.  
AC Q53693;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE Leader peptide.  
OS Streptomyces avermitilis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=33903;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CU-18;  
RX MEDLINE=96105213; PubMed=8529876;  
RA De Rossi E., Leva R., Gusberti L., Manachini P.L., Riccardi G.;  
RT "Cloning, sequencing and expression of the ilvBNC gene cluster from  
RT Streptomyces avermitilis.";  
RL Gene 166:127-132(1995).  
DR EMBL; L39268; AAA93097.1; -.  
SQ SEQUENCE 13 AA; 1499 MW; 5E080E40FD4E32C7 CRC64;

Query Match 37.5%; Score 3; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TRI 4  
|||  
Db 3 TRI 5

RESULT 11

Q9NR93

ID Q9NR93 PRELIMINARY; PRT; 13 AA.  
AC Q9NR93;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE ULIP (Fragment).  
GN ULIP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20287517; PubMed=10748015;  
 RA Matsuo T., Stauffer J.K., Walker R.L., Meltzer P., Thiele C.J.;  
 RT "Structure and promoter analysis of the human unc-33-like  
 phosphoprotein gene. E-box required for maximal expression in  
 neuroblastoma and myoblasts.";  
 RL J. Biol. Chem. 275:16560-16568(2000).  
 DR EMBL; AF246692; AAF76324.1; -.  
 FT NON\_TER 13 13  
 SQ SEQUENCE 13 AA; 1536 MW; 517F5F2CE06D4333 CRC64;

Query Match 37.5%; Score 3; DB 4; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 RIT 5  
 |||  
 Db 11 RIT 13

# RESULT 12

## Q8SAT3

ID Q8SAT3 PRELIMINARY; PRT; 13 AA.  
 AC Q8SAT3;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Non-C4 chloroplastic NADP-malic enzyme (Fragment).  
 GN CHLME2-2.  
 OS Flaveria robusta.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; campanulids; Asterales; Asteraceae; Asteroideae; Tageteae;  
 OC Flaveria.  
 OX NCBI\_TaxID=163089;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21648948; PubMed=11788758;  
 RA Lai L.B., Wang L., Nelson T.M.;  
 RT "Distinct But Conserved Functions for Two Chloroplastic NADP-Malic  
 Enzyme Isoforms in C(3) and C(4) Flaveria Species.";  
 RL Plant Physiol. 128:125-139(2002).  
 DR EMBL; AF288914; AAL74058.1; -.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 13 AA; 1697 MW; 95393F9D91422775 CRC64;

Query Match 37.5%; Score 3; DB 10; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ESC 8  
 |||  
 Db 1 ESC 3

# RESULT 13

Q9S922

ID Q9S922 PRELIMINARY; PRT; 13 AA.  
AC Q9S922;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE S3 SGP=S-locus specific glycoprotein (Fragment).  
OS Brassica oleracea (Cauliflower).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Brassica.  
OX NCBI\_TaxID=3712;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92090397; PubMed=1752245;  
RA Gaude T., Denoroy L., Dumas C.;  
RT "Use of a fast protein electrophoretic purification procedure for N-  
RT terminal sequence analysis to identify S-locus related proteins in  
RT stigmas of Brassica oleracea.";  
RL Electrophoresis 12:646-653(1991).  
FT NON\_TER 1 1  
FT NON\_TER 13 13  
SQ SEQUENCE 13 AA; 1351 MW; 9C23EDFF999601B5 CRC64;

Query Match 37.5%; Score 3; DB 10; Length 13;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TES 7  
|||  
Db 7 TES 9

RESULT 14

Q8SAT4

ID Q8SAT4 PRELIMINARY; PRT; 13 AA.  
AC Q8SAT4;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Non-C4 chloroplastic NADP-malic enzyme (Fragment).  
GN CHLME2.  
OS Flaveria angustifolia.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; campanulids; Asterales; Asteraceae; Asteroideae; Tageteae;  
OC Flaveria.  
OX NCBI\_TaxID=163088;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21648948; PubMed=11788758;  
RA Lai L.B., Wang L., Nelson T.M.;  
RT "Distinct But Conserved Functions for Two Chloroplastic NADP-Malic  
RT Enzyme Isoforms in C(3) and C(4) Flaveria Species.";  
RL Plant Physiol. 128:125-139(2002).  
DR EMBL; AF288896; AAL74057.1; -.  
FT NON\_TER 1 1

SQ SEQUENCE 13 AA; 1697 MW; 95393F9D91422775 CRC64;

Query Match 37.5%; Score 3; DB 10; Length 13;  
Best Local Similarity 100.0%; Pred. No. 3.6e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ESC 8  
|||  
Db 1 ESC 3

RESULT 15

Q9PRM3

ID Q9PRM3 PRELIMINARY; PRT; 15 AA.  
AC Q9PRM3;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE 17 kDa major immunophilin (Fragment).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=94072550; PubMed=7504525;  
RA Yem A.W., Reardon I.M., Leone J.W., Heinrikson R.L., Deibel M.R.Jr.;  
RT "An active FK506-binding domain of 17,000 daltons is isolated  
RT following limited proteolysis of chicken thymus hsp56."  
RL Biochemistry 32:12571-12576(1993).  
SQ SEQUENCE 15 AA; 1606 MW; 6269732398D1B71C CRC64;

Query Match 37.5%; Score 3; DB 13; Length 15;  
Best Local Similarity 100.0%; Pred. No. 4e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TES 7  
|||  
Db 7 TES 9

Search completed: November 13, 2003, 10:38:18  
Job time : 20.6667 secs



OM protein - protein search, using sw model

Run on: November 13, 2003, 10:27:36 ; Search time 9 Seconds  
(without alignments)  
37.610 Million cell updates/sec

Title: US-09-228-866-9  
Perfect score: 8  
Sequence: 1 CTRITESC 8

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 118358

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	8	100.0	8	1	US-08-526-710-9	Sequence 9, Appli
2	8	100.0	8	3	US-08-862-855-9	Sequence 9, Appli
3	8	100.0	8	3	US-09-226-985-9	Sequence 9, Appli
4	8	100.0	8	4	US-09-227-906-9	Sequence 9, Appli
5	4	50.0	13	6	5245013-7	Patent No. 5245013
6	4	50.0	14	1	US-08-205-719-5	Sequence 5, Appli
7	4	50.0	15	2	US-08-310-912A-38	Sequence 38, Appl
8	4	50.0	15	3	US-08-841-089-38	Sequence 38, Appl
9	4	50.0	15	3	US-09-301-085-38	Sequence 38, Appl
10	4	50.0	15	5	PCT-US95-04570-38	Sequence 38, Appl
11	4	50.0	15	5	PCT-US95-04589-38	Sequence 38, Appl

12	4	50.0	16	3	US-08-602-999A-203	Sequence 203, App
13	4	50.0	16	4	US-09-500-124-203	Sequence 203, App
14	4	50.0	18	1	US-08-478-312-13	Sequence 13, Appl
15	4	50.0	18	1	US-08-478-312-38	Sequence 38, Appl
16	4	50.0	18	1	US-08-478-312-41	Sequence 41, Appl
17	4	50.0	18	1	US-08-478-312-42	Sequence 42, Appl
18	4	50.0	18	1	US-08-478-312-44	Sequence 44, Appl
19	4	50.0	18	1	US-08-478-312-45	Sequence 45, Appl
20	4	50.0	18	1	US-08-485-302-13	Sequence 13, Appl
21	4	50.0	18	1	US-08-485-302-38	Sequence 38, Appl
22	4	50.0	18	1	US-08-485-302-41	Sequence 41, Appl
23	4	50.0	18	1	US-08-485-302-42	Sequence 42, Appl
24	4	50.0	18	1	US-08-485-302-44	Sequence 44, Appl
25	4	50.0	18	1	US-08-485-302-45	Sequence 45, Appl
26	4	50.0	20	1	US-07-994-277A-11	Sequence 11, Appl
27	4	50.0	21	3	US-08-468-011A-22	Sequence 22, Appl
28	4	50.0	21	4	US-09-236-468A-22	Sequence 22, Appl
29	3	37.5	7	1	US-08-248-538-14	Sequence 14, Appl
30	3	37.5	7	1	US-08-539-798-10	Sequence 10, Appl
31	3	37.5	7	1	US-08-329-560-10	Sequence 10, Appl
32	3	37.5	7	1	US-08-690-605-15	Sequence 15, Appl
33	3	37.5	7	2	US-08-310-912A-90	Sequence 90, Appl
34	3	37.5	7	3	US-09-045-632-44	Sequence 44, Appl
35	3	37.5	7	3	US-08-804-439A-75	Sequence 75, Appl
36	3	37.5	7	3	US-08-720-229-75	Sequence 75, Appl
37	3	37.5	7	3	US-08-893-526A-6	Sequence 6, Appli
38	3	37.5	7	3	US-08-841-089-90	Sequence 90, Appl
39	3	37.5	7	3	US-09-147-933-31	Sequence 31, Appl
40	3	37.5	7	3	US-09-301-085-90	Sequence 90, Appl
41	3	37.5	7	4	US-08-775-703A-14	Sequence 14, Appl
42	3	37.5	7	5	PCT-US95-04570-90	Sequence 90, Appl
43	3	37.5	7	5	PCT-US95-04589-90	Sequence 90, Appl
44	3	37.5	8	1	US-08-362-453-6	Sequence 6, Appli
45	3	37.5	8	1	US-08-259-550A-30	Sequence 30, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-526-710-9

; Sequence 9, Application US/08526710

; Patent No. 5622699

; GENERAL INFORMATION:

; APPLICANT: Ruoslahti, Erkki

; APPLICANT: Pasqualini, Renata

; TITLE OF INVENTION: Method of Identifying Molecules That

; TITLE OF INVENTION: Home to a Selected Organ In Vivo

; NUMBER OF SEQUENCES: 44

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Campbell and Flores

; STREET: 4370 La Jolla Village Drive, Suite 700

; CITY: San Diego

; STATE: California

; COUNTRY: United States

; ZIP: 92122

; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/526,710
; FILING DATE: 11-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-526-710-9

```

```

Query Match          100.0%; Score 8; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 CTRITESC 8
        |||||
Db      1 CTRITESC 8

```

# RESULT 2

US-08-862-855-9

```

; Sequence 9, Application US/08862855
; Patent No. 6068829
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,855
; FILING DATE:

```

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-862-855-9

```

```

Query Match          100.0%; Score 8; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 CTRITESC 8
        |||||
Db      1 CTRITESC 8

```

```

RESULT 3
US-09-226-985-9
; Sequence 9, Application US/09226985
; Patent No. 6296832
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Molecules That Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/226,985
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,855
; FILING DATE: 23-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 3423
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-226-985-9

```

```

Query Match          100.0%; Score 8; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      8; Conservative    0; Mismatches    0; Indels      0; Gaps      0;

```

```

Qy      1 CTRITESC 8
        |||||
Db      1 CTRITESC 8

```

#### RESULT 4

US-09-227-906-9

```

; Sequence 9, Application US/09227906
; Patent No. 6306365
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/227,906
; FILING DATE:

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,855
; FILING DATE: 23-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 3424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-227-906-9

```

```

Query Match          100.0%; Score 8; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 CTRITESC 8
        |||||
Db      1 CTRITESC 8

```

```

RESULT 5
5245013-7
;Patent No. 5245013
; APPLICANT: Ulevitch, Richard;Tobias, Peter
; TITLE OF INVENTION: ACUTE PHASE PROTEIN MODULATING ENDOTOXIC
;ACTIVITY OF LIPOPOLYSACCHARIDES, ASSAY METHODS AND POLYPEPTIDES
; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/367,454
; FILING DATE: 01-JUN-1989
; APPLICATION NUMBER: 6,710
; FILING DATE: 30-DEC-1986
; APPLICATION NUMBER: 728,833
; FILING DATE: 30-APR-1985
;SEQ ID NO:7:
; LENGTH: 13
5245013-7

```

```

Query Match          50.0%; Score 4; DB 6; Length 13;
Best Local Similarity 100.0%; Pred. No. 40;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      2 TRIT 5

```

Db           ||||  
              7 TRIT 10

RESULT 6  
US-08-205-719-5  
; Sequence 5, Application US/08205719  
; Patent No. 5705398  
; GENERAL INFORMATION:  
; APPLICANT: Mintz, D. N.  
; APPLICANT: Tobias, P. S.  
; APPLICANT: Ulevitch, R. J.  
; TITLE OF INVENTION: SYSTEM AND METHODS FOR IDENTIFYING  
; TITLE OF INVENTION: THERAPEUTIC INHIBITORS OF LPS-MEDIATED SEPSIS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute, Office of  
; ADDRESSEE: Patent Counsel  
; STREET: 10666 No. 5705398th Torrey Pines Road, TPC 8  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/205,719  
; FILING DATE: 02-MAR-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitting, Thomas  
; REGISTRATION NUMBER: 34,163  
; REFERENCE/DOCKET NUMBER: SCR1201P; TSRI324.0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-554-2937  
; TELEFAX: 619-554-6312  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: N-terminal  
US-08-205-719-5

Query Match           50.0%; Score 4; DB 1; Length 14;  
Best Local Similarity   100.0%; Pred. No. 42;  
Matches       4; Conservative   0; Mismatches   0; Indels       0; Gaps       0;

Qy           2 TRIT 5  
              ||||  
Db           7 TRIT 10

RESULT 7  
 US-08-310-912A-38  
 ; Sequence 38, Application US/08310912A  
 ; Patent No. 5981730  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ausubel, Frederick M.  
 ; APPLICANT: Staskawicz, Brian J.  
 ; APPLICANT: Brent, Andrew F.  
 ; APPLICANT: Dahlbeck, Douglas  
 ; APPLICANT: Katagiri, Fumiaki  
 ; APPLICANT: Kunkel, Barbara N.  
 ; APPLICANT: Mindrinos, Michael N.  
 ; APPLICANT: Yu, Guo-Liang  
 ; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND DETECTION  
 ; TITLE OF INVENTION: METHODS  
 ; NUMBER OF SEQUENCES: 208  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson P.C.  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02110-2904  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/310,912A  
 ; FILING DATE: September 22, 1994  
 ; CLASSIFICATION: 536  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/227,360  
 ; FILING DATE: April 13, 1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Lech, Karen F.  
 ; REGISTRATION NUMBER: 35,238  
 ; REFERENCE/DOCKET NUMBER: 00786/254001  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 542-5070  
 ; TELEFAX: (617) 542-8906  
 ; TELEX: 100254  
 ; INFORMATION FOR SEQ ID NO: 38:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 15 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: not relevant  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-310-912A-38

Query Match 50.0%; Score 4; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITES 7



Db           ||||  
              1 ITES 4

RESULT 8  
US-08-841-089-38  
; Sequence 38, Application US/08841089  
; Patent No. 6127607  
; GENERAL INFORMATION:  
; APPLICANT: Ausubel, Frederick M.  
; APPLICANT: Staskawicz, Brian J.  
; APPLICANT: Brent, Andrew F.  
; APPLICANT: Dahlbeck, Douglas  
; APPLICANT: Katagiri, Fumiaki  
; APPLICANT: Kunkel, Barbara N.  
; APPLICANT: Mindrinos, Michael N.  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: RPS2 GENE AND USES THEREOF  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street Suite 3100  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2904  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/841,089  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/227,360  
; FILING DATE: 13-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 00786/230001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 100254  
; INFORMATION FOR SEQ ID NO: 38:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 15 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-841-089-38

Query Match                   50.0%; Score 4; DB 3; Length 15;  
Best Local Similarity       100.0%; Pred. No. 45;

```

Matches      4;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

Qy           4 ITES 7
            ||||
Db           1 ITES 4

```

RESULT 9

US-09-301-085-38

```

; Sequence 38, Application US/09301085
; Patent No. 6262248
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Staskawicz, Brian J.
; APPLICANT: Brent, Andrew F.
; APPLICANT: Dahlbeck, Douglas
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kunkel, Barbara N.
; APPLICANT: Mindrinos, Michael N.
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND
; TITLE OF INVENTION: DETECTION METHODS
; FILE REFERENCE: 00786/254002
; CURRENT APPLICATION NUMBER: US/09/301,085
; CURRENT FILING DATE: 1999-04-28
; EARLIER APPLICATION NUMBER: 08/310,912
; EARLIER FILING DATE: 1994-09-22
; EARLIER APPLICATION NUMBER: 08/227,360
; EARLIER FILING DATE: 1994-04-13
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-301-085-38

```

```

Query Match          50.0%;  Score 4;  DB 3;  Length 15;
Best Local Similarity 100.0%;  Pred. No. 45;
Matches      4;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

Qy           4 ITES 7
            ||||
Db           1 ITES 4

```

RESULT 10

PCT-US95-04570-38

```

; Sequence 38, Application PC/TUS9504570
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Staskawicz, Brian J.
; APPLICANT: Brent, Andrew F.
; APPLICANT: Dahlbeck, Douglas
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kunkel, Barbara N.
; APPLICANT: Mindrinos, Michael N.

```

```

; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: RPS2 GENE AND USES THEREOF
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street Suite 3100
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2904
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04570
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,360
; FILING DATE: 13-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/230001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 100254
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-04570-38

```

```

Query Match          50.0%; Score 4; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 45;
Matches    4; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

```

```

Qy          4 ITES 7
            ||||
Db          1 ITES 4

```

```

RESULT 11
PCT-US95-04589-38
; Sequence 38, Application PC/TUS9504589
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Staskawicz, Brian J.
; APPLICANT: Brent, Andrew F.
; APPLICANT: Dahlbeck, Douglas
; APPLICANT: Katagiri, Fumiaki

```

```

; APPLICANT: Kunkel, Barbara N.
; APPLICANT: Mindrinos, Michael N.
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: RPS2 GENE AND USES THEREOF
; NUMBER OF SEQUENCES: 201
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street Suite 3100
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2904
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04589
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,360
; FILING DATE: 13-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/230001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 100254
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-04589-38

```

```

Query Match          50.0%; Score 4; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 45;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      4 ITES 7
      ||||
Db      1 ITES 4

```

```

RESULT 12
US-08-602-999A-203
; Sequence 203, Application US/08602999A
; Patent No. 6184205
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.

```

```

; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,999A
; FILING DATE: 16-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 203:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-602-999A-203

```

```

Query Match          50.0%; Score 4; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 48;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      2 TRIT 5
        ||||
Db      13 TRIT 16

```

```

RESULT 13
US-09-500-124-203
; Sequence 203, Application US/09500124
; Patent No. 6432920
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.

```

```

; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/500,124
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/602,999
; FILING DATE: 16-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 203:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-09-500-124-203

```

```

Query Match          50.0%; Score 4; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 48;
Matches    4; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

```

```

Qy          2 TRIT 5
           ||||
Db          13 TRIT 16

```

```

RESULT 14
US-08-478-312-13
; Sequence 13, Application US/08478312
; Patent No. 5654276
; GENERAL INFORMATION:
; APPLICANT: Barrett, Ronald W.

```

```

; APPLICANT: England, Bruce
; APPLICANT: Schatz, Peter
; APPLICANT: Sloan, Derek
; APPLICANT: Chen, Min-Jia
; TITLE OF INVENTION: Peptides and Compounds That Bind to the IL-5
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Affymax Technologies, N.V.
; STREET: 4001 Miranda Ave.
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,312
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens, Lauren L.
; REGISTRATION NUMBER: 36,691
; REFERENCE/DOCKET NUMBER: 1088.1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-496-2300
; TELEFAX: 415-424-0832
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-478-312-13

```

```

Query Match          50.0%; Score 4; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 54;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 CTRI 4
        ||||
Db      4 CTRI 7

```

```

RESULT 15
US-08-478-312-38
; Sequence 38, Application US/08478312
; Patent No. 5654276
; GENERAL INFORMATION:
; APPLICANT: Barrett, Ronald W.
; APPLICANT: England, Bruce
; APPLICANT: Schatz, Peter
; APPLICANT: Sloan, Derek

```

```

; APPLICANT: Chen, Min-Jia
; TITLE OF INVENTION: Peptides and Compounds That Bind to the IL-5
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Affymax Technologies, N.V.
; STREET: 4001 Miranda Ave.
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,312
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens, Lauren L.
; REGISTRATION NUMBER: 36,691
; REFERENCE/DOCKET NUMBER: 1088.1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-496-2300
; TELEFAX: 415-424-0832
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-478-312-38

```

```

Query Match          50.0%; Score 4; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 54;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 CTRI 4
        ||||
Db      4 CTRI 7

```

```

Search completed: November 13, 2003, 10:41:57
Job time : 10 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 10:27:36 ; Search time 23.625 Seconds  
(without alignments)  
37.610 Million cell updates/sec

Title: US-09-228-866-16  
Perfect score: 21  
Sequence: 1 WRCVLREGPAGGCAWFNRHRL 21

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 118358

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	21	100.0	21	1	US-08-526-710-16 Sequence 16, Appl
2	21	100.0	21	3	US-08-862-855-16 Sequence 16, Appl
3	21	100.0	21	3	US-09-226-985-16 Sequence 16, Appl
4	21	100.0	21	4	US-09-227-906-16 Sequence 16, Appl
5	9	42.9	9	1	US-08-526-710-20 Sequence 20, Appl
6	9	42.9	9	3	US-08-862-855-20 Sequence 20, Appl
7	9	42.9	9	3	US-09-226-985-20 Sequence 20, Appl
8	9	42.9	9	4	US-09-227-906-20 Sequence 20, Appl
9	5	23.8	8	3	US-08-159-339A-864 Sequence 864, App
10	5	23.8	9	3	US-08-159-339A-280 Sequence 280, App
11	5	23.8	10	3	US-08-159-339A-301 Sequence 301, App

12	5	23.8	12	2	US-08-764-640-85	Sequence 85, Appl
13	5	23.8	12	3	US-08-973-225-85	Sequence 85, Appl
14	5	23.8	12	3	US-09-244-298A-85	Sequence 85, Appl
15	5	23.8	12	3	US-09-516-704-85	Sequence 85, Appl
16	5	23.8	12	4	US-09-549-090-85	Sequence 85, Appl
17	5	23.8	12	4	US-09-832-230A-85	Sequence 85, Appl
18	5	23.8	13	2	US-08-572-951-19	Sequence 19, Appl
19	5	23.8	13	2	US-08-572-951-20	Sequence 20, Appl
20	5	23.8	13	2	US-08-572-951-40	Sequence 40, Appl
21	5	23.8	13	3	US-09-258-754-113	Sequence 113, App
22	5	23.8	13	3	US-09-042-107-113	Sequence 113, App
23	5	23.8	20	4	US-09-392-055-6	Sequence 6, Appli
24	5	23.8	21	4	US-09-392-055-12	Sequence 12, Appl
25	4	19.0	7	1	US-07-989-962-8	Sequence 8, Appli
26	4	19.0	7	1	US-07-989-962-15	Sequence 15, Appl
27	4	19.0	7	1	US-07-989-962-16	Sequence 16, Appl
28	4	19.0	7	1	US-07-989-962-21	Sequence 21, Appl
29	4	19.0	7	1	US-07-989-962-22	Sequence 22, Appl
30	4	19.0	7	1	US-08-213-402-8	Sequence 8, Appli
31	4	19.0	7	1	US-08-213-402-15	Sequence 15, Appl
32	4	19.0	7	1	US-08-213-402-16	Sequence 16, Appl
33	4	19.0	7	1	US-08-213-402-21	Sequence 21, Appl
34	4	19.0	7	1	US-08-213-402-22	Sequence 22, Appl
35	4	19.0	7	1	US-08-459-888-8	Sequence 8, Appli
36	4	19.0	7	1	US-08-459-888-15	Sequence 15, Appl
37	4	19.0	7	1	US-08-459-888-16	Sequence 16, Appl
38	4	19.0	7	1	US-08-459-888-21	Sequence 21, Appl
39	4	19.0	7	1	US-08-459-888-22	Sequence 22, Appl
40	4	19.0	7	2	US-08-460-452-8	Sequence 8, Appli
41	4	19.0	7	2	US-08-460-452-15	Sequence 15, Appl
42	4	19.0	7	2	US-08-460-452-16	Sequence 16, Appl
43	4	19.0	7	2	US-08-460-452-21	Sequence 21, Appl
44	4	19.0	7	2	US-08-460-452-22	Sequence 22, Appl
45	4	19.0	7	2	US-08-459-557-8	Sequence 8, Appli

# ALIGNMENTS

## RESULT 1

US-08-526-710-16

; Sequence 16, Application US/08526710

; Patent No. 5622699

; GENERAL INFORMATION:

; APPLICANT: Ruoslahti, Erkki

; APPLICANT: Pasqualini, Renata

; TITLE OF INVENTION: Method of Identifying Molecules That

; TITLE OF INVENTION: Home to a Selected Organ In Vivo

; NUMBER OF SEQUENCES: 44

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Campbell and Flores

; STREET: 4370 La Jolla Village Drive, Suite 700

; CITY: San Diego

; STATE: California

; COUNTRY: United States

; ZIP: 92122

; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/526,710
; FILING DATE: 11-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-526-710-16

```

```

Query Match          100.0%; Score 21; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.2e-15;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 WRCVLREGPAGGCAWFNRHRL 21
        ||||||||||||||||
Db      1 WRCVLREGPAGGCAWFNRHRL 21

```

```

RESULT 2
US-08-862-855-16
; Sequence 16, Application US/08862855
; Patent No. 6068829
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,855
; FILING DATE:

```

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-862-855-16

```

```

Query Match          100.0%; Score 21; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.2e-15;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 WRCVLREGPAGGCAWFNRHRL 21
         ||||||||||||||||||||
Db      1 WRCVLREGPAGGCAWFNRHRL 21

```

```

RESULT 3
US-09-226-985-16
; Sequence 16, Application US/09226985
; Patent No. 6296832
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Molecules That Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/226,985
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,855
; FILING DATE: 23-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 3423
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-226-985-16

```

```

Query Match          100.0%; Score 21; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.2e-15;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 WRCVLREGPAGGCAWFNRHRL 21
        ||||||||||||||||
Db      1 WRCVLREGPAGGCAWFNRHRL 21

```

#### RESULT 4

```

US-09-227-906-16
; Sequence 16, Application US/09227906
; Patent No. 6306365
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/227,906
; FILING DATE:

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,855
; FILING DATE: 23-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 3424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-227-906-16

```

```

Query Match          100.0%; Score 21; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.2e-15;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 WRCVLREGPAGGCAWFNRHRL 21
        |||||
Db      1 WRCVLREGPAGGCAWFNRHRL 21

```

```

RESULT 5
US-08-526-710-20
; Sequence 20, Application US/08526710
; Patent No. 5622699
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/526,710  
 ; FILING DATE: 11-SEP-1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Campbell, Cathryn A.  
 ; REGISTRATION NUMBER: 31,815  
 ; REFERENCE/DOCKET NUMBER: P-LJ 1779  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (619) 535-9001  
 ; TELEFAX: (619) 535-8949  
 ; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 9 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 US-08-526-710-20

Query Match 42.9%; Score 9; DB 1; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VLREGPAGG 12  
 |||||  
 Db 1 VLREGPAGG 9

RESULT 6

US-08-862-855-20

; Sequence 20, Application US/08862855  
 ; Patent No. 6068829  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ruoslahti, Erkki  
 ; APPLICANT: Pasqualini, Renata  
 ; TITLE OF INVENTION: Method of Identifying Molecules That  
 ; TITLE OF INVENTION: Home to a Selected Organ In Vivo  
 ; NUMBER OF SEQUENCES: 44  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Campbell & Flores LLP  
 ; STREET: 4370 La Jolla Village Drive, Suite 700  
 ; CITY: San Diego  
 ; STATE: California  
 ; COUNTRY: United States  
 ; ZIP: 92122  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/862,855  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/526,710  
 ; FILING DATE: 11-SEP-1995  
 ; PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-862-855-20

```

```

Query Match          42.9%; Score 9; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      4 VLREGPAGG 12
        |||||
Db      1 VLREGPAGG 9

```

# RESULT 7

```

US-09-226-985-20
; Sequence 20, Application US/09226985
; Patent No. 6296832
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Molecules That Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/226,985
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997

```



```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,855
; FILING DATE: 23-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 3423
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-226-985-20

```

```

Query Match          42.9%; Score 9; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      4 VLREGPAGG 12
        |||||
Db      1 VLREGPAGG 9

```

# RESULT 8

```

US-09-227-906-20
; Sequence 20, Application US/09227906
; Patent No. 6306365
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Method of Identifying Molecules That
; TITLE OF INVENTION: Home to a Selected Organ In Vivo
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/227,906
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,710
; FILING DATE: 11-SEP-1995
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/813,273
; FILING DATE: 10-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/862,855
; FILING DATE: 23-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 3424
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-227-906-20

```

```

Query Match          42.9%; Score 9; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      4 VLREGPAGG 12
        |||||
Db      1 VLREGPAGG 9

```

# RESULT 9

```

US-08-159-339A-864
; Sequence 864, Application US/08159339A
; Patent No. 6037135
; GENERAL INFORMATION:
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Sette, Alessandro
; APPLICANT: Celis, Esteban
; TITLE OF INVENTION: HLA Binding peptides and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 1254
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,339A
; FILING DATE: 29-NOV-1993
; CLASSIFICATION: 424

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,666
; FILING DATE: 07-AUG-1992
; APPLICATION NUMBER: US 08/027,746
; FILING DATE: 05-MAR-1993
; APPLICATION NUMBER: US 08/103,396
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 018623-005030US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX:
; INFORMATION FOR SEQ ID NO: 864:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-159-339A-864

```

```

Query Match          23.8%; Score 5; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      8 GPAGG 12
        |||||
Db      1 GPAGG 5

```

```

RESULT 10
US-08-159-339A-280
; Sequence 280, Application US/08159339A
; Patent No. 6037135
; GENERAL INFORMATION:
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Sette, Alessandro
; APPLICANT: Celis, Esteban
; TITLE OF INVENTION: HLA Binding peptides and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 1254
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,339A
; FILING DATE: 29-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,666
; FILING DATE: 07-AUG-1992
; APPLICATION NUMBER: US 08/027,746
; FILING DATE: 05-MAR-1993
; APPLICATION NUMBER: US 08/103,396
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 018623-005030US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX:
; INFORMATION FOR SEQ ID NO: 280:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-159-339A-280

```

```

Query Match          23.8%; Score 5; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      10 AGGCA 14
        |||||
Db      4 AGGCA 8

```

```

RESULT 11
US-08-159-339A-301
; Sequence 301, Application US/08159339A
; Patent No. 6037135
; GENERAL INFORMATION:
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Sette, Alessandro
; APPLICANT: Celis, Esteban
; TITLE OF INVENTION: HLA Binding peptides and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 1254
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,339A
; FILING DATE: 29-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,666
; FILING DATE: 07-AUG-1992
; APPLICATION NUMBER: US 08/027,746
; FILING DATE: 05-MAR-1993
; APPLICATION NUMBER: US 08/103,396
; FILING DATE: 06-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 018623-005030US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX:
; INFORMATION FOR SEQ ID NO: 301:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-159-339A-301

```

```

Query Match          23.8%; Score 5; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 32;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      10 AGGCA 14
        |||||
Db      5 AGGCA 9

```

# RESULT 12

US-08-764-640-85

```

; Sequence 85, Application US/08764640
; Patent No. 5869451
; Patent No. 5869451 5837683
; GENERAL INFORMATION:
; APPLICANT: Dower, William J.
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Cwirla, Steven E.
; APPLICANT: Gates, Christian
; APPLICANT: Schatz, Peter J.
; APPLICANT: Balasubramanian, Palaniappan
; APPLICANT: Wagstrom, Christopher R.
; APPLICANT: Hendren, Richard W.
; APPLICANT: Deprince, Randolph B.
; APPLICANT: Podduturi, Surekha

```

```

; APPLICANT: Yin, Qun
; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
; TITLE OF INVENTION: RECEPTOR
; NUMBER OF SEQUENCES: 244
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Glaxo Wellcome
; STREET: Five Moore Drive, P.O. Box 13398
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/764,640
; FILING DATE: 11-DEC-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hrubiec, Robert T.
; REGISTRATION NUMBER: 36,392
; REFERENCE/DOCKET NUMBER: PK3281
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-248-1000
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-764-640-85

```

```

Query Match          23.8%; Score 5; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 37;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      3 CVLRE 7
        |||||
Db      1 CVLRE 5

```

# RESULT 13

US-08-973-225-85

```

; Sequence 85, Application US/08973225A
; Patent No. 6083913
; GENERAL INFORMATION:
; APPLICANT: Dower, William J.
;            Barrett, Ronald W.
;            Cwirla, Steven E.
;            Duffin, David J.
;            Gates, Christian
;            Haselden, Sherril S.
;            Mattheakis, Larry C.
;            Schatz, Peter J.

```

```

;           Wagstrom, Christopher R.
;           Wrighton, Nicholas C.
;   TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
;                       THROMBOPOIETIN RECEPTOR
;   NUMBER OF SEQUENCES: 232
;   CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Glaxo Wellcome
;       STREET: Five Moore Drive, P.O. Box 13398
;       CITY: Research Triangle Park
;       STATE: NC
;       COUNTRY: USA
;       ZIP: 27709
;   COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/973,225A
;       FILING DATE: 04-Dec-1997
;   ATTORNEY/AGENT INFORMATION:
;       NAME: Hrubiec, Robert T.
;       REGISTRATION NUMBER: 36,392
;       REFERENCE/DOCKET NUMBER: PK3065USW
;   TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 919-248-1000
;   INFORMATION FOR SEQ ID NO: 85:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 12 amino acids
;           TYPE: amino acid
;           STRANDEDNESS: <Unknown>
;           TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 85:
US-08-973-225-85

```

```

Query Match          23.8%;  Score 5;  DB 3;  Length 12;
Best Local Similarity 100.0%;  Pred. No. 37;
Matches      5;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

```

```

Qy      3  CVLRE 7
        |||||
Db      1  CVLRE 5

```

```

RESULT 14
US-09-244-298A-85
; Sequence 85, Application US/09244298A
; Patent No. 6121238
;   GENERAL INFORMATION:
;   APPLICANT: Dower, William J.
;   APPLICANT: Barrett, Ronald W.
;   APPLICANT: Cwirla, Steven E.
;   APPLICANT: Gates, Christian
;   APPLICANT: Schatz, Peter J.
;   APPLICANT: Balasubramanian, Palaniappan
;   APPLICANT: Wagstrom, Christopher R.

```

```

; APPLICANT: Hendren, Richard W.
; APPLICANT: Deprince, Randolph B.
; APPLICANT: Podduturi, Surekha
; APPLICANT: Yin, Qun
; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
; TITLE OF INVENTION: RECEPTOR
; NUMBER OF SEQUENCES: 244
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Glaxo Wellcome
; STREET: Five Moore Drive, P.O. Box 13398
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/244,298A
; FILING DATE: 11-DEC-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hrubiec, Robert T.
; REGISTRATION NUMBER: 36,392
; REFERENCE/DOCKET NUMBER: PK3281
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-248-1000
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-244-298A-85

```

```

Query Match          23.8%; Score 5; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 37;
Matches      5; Conservative    0; Mismatches    0; Indels      0; Gaps      0;

```

```

Qy          3 CVLRE 7
            |||||
Db          1 CVLRE 5

```

```

RESULT 15
US-09-516-704-85
; Sequence 85, Application US/09516704
; Patent No. 6251864
; GENERAL INFORMATION:
; APPLICANT: Dower, William J.
;            Barrett, Ronald W.
;            Cwirla, Steven E.
;            Gates, Christian
;            Schatz, Peter J.

```



```

;           Balasubramanian, Palaniappan
;           Wagstrom, Christopher R.
;           Hendren, Richard W.
;           Deprince, Randolph B.
;           Podduturi, Surekha
;   TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
;                       RECEPTOR
;   NUMBER OF SEQUENCES: 244
;   CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Glaxo Wellcome
;       STREET: Five Moore Drive, P.O. Box 13398
;       CITY: Research Triangle Park
;       STATE: NC
;       COUNTRY: USA
;       ZIP: 27709
;   COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/09/516,704
;       FILING DATE: 01-Mar-2000
;       CLASSIFICATION: <Unknown>
;   ATTORNEY/AGENT INFORMATION:
;       NAME: Hrubiec, Robert T.
;       REGISTRATION NUMBER: 36,392
;       REFERENCE/DOCKET NUMBER: PK3281
;   TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 919-248-1000
;   INFORMATION FOR SEQ ID NO: 85:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 12 amino acids
;           TYPE: amino acid
;           STRANDEDNESS: <Unknown>
;           TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       SEQUENCE DESCRIPTION: SEQ ID NO: 85:
US-09-516-704-85

Query Match          23.8%; Score 5; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 37;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Qy           3 CVLRE 7
             |||||
Db           1 CVLRE 5

```

Search completed: November 13, 2003, 10:41:57  
Job time : 23.625 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:53:11 ; Search time 67.375 Seconds  
(without alignments)  
49.473 Million cell updates/sec

Title: US-09-228-866-16  
Perfect score: 21  
Sequence: 1 WRCVLREGPAGGCAWFNRHRL 21

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 358712

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	21	100.0	21	18	AAW13415	Brain homing pepti	
2	21	100.0	21	21	AAB12002	Brain homing pepti	
3	21	100.0	21	22	AAE11808	Phage peptide #16	
4	21	100.0	21	23	AAU10719	Brain homing pepti	
5	21	100.0	21	24	ABU59531	Brain receptor tar	
6	9	42.9	9	18	AAW13435	Brain homing pepti	
7	9	42.9	9	21	AAB12006	Brain homing pepti	
8	9	42.9	9	22	AAE11812	Phage peptide #20	
9	9	42.9	9	23	AAU10723	Brain homing pepti	
10	5	23.8	9	15	AAV37974	Human cERB2 oncoge	
11	5	23.8	9	20	AAV45544	Immunogenic peptid	
12	5	23.8	9	22	AAG88625	HER2/NEU DR superm	
13	5	23.8	9	22	AAG88807	HER2/neu epitope H	
14	5	23.8	10	15	AAV37995	Human cERB2 oncoge	
15	5	23.8	10	20	AAV45565	Immunogenic peptid	
16	5	23.8	10	22	AAG97585	Human complementar	
17	5	23.8	10	22	AAG88806	HER2/neu epitope H	
18	5	23.8	10	23	AAU71491	Human MHC molecule	
19	5	23.8	10	23	AAU71652	Human MHC class I	
20	5	23.8	10	23	AAU71769	Human MHC molecule	
21	5	23.8	12	18	AAW36680	Thrombopoietin rec	
22	5	23.8	12	18	AAW09529	Thrombopoietin rec	
23	5	23.8	12	22	AAU25899	Human thrombopoiet	
24	5	23.8	13	18	AAW18067	Soluble starch syn	
25	5	23.8	13	20	AAV48704	Membrane dipeptida	
26	5	23.8	14	22	AAM00404	Human protein frag	
27	5	23.8	14	22	AAG84639	MAGE3 DR supermoti	
28	5	23.8	15	22	AAG88316	HER2/NEU DR superm	
29	5	23.8	15	22	AAG88626	HER2/NEU DR superm	
30	5	23.8	15	23	ABG66107	IgE Fcepsilon RI b	
31	5	23.8	18	21	AAB34029	Human secreted pro	
32	5	23.8	18	22	AAE10675	Human melanoma ass	
33	5	23.8	18	23	AAU90893	Insulin/insulin-li	
34	5	23.8	18	23	AAU85064	Human MAGE-3 segme	
35	4	19.0	7	20	AAV14807	Cadherin superfami	
36	4	19.0	7	23	ABG77639	Targetting peptide	
37	4	19.0	7	23	ABJ04527	Molt-4 leukaemia c	
38	4	19.0	7	24	ABG73798	Molecular library-	
39	4	19.0	7	24	ABP74898	Proteome analysis	
40	4	19.0	8	15	AAR60435	Antiproliferative	
41	4	19.0	8	18	AAW13414	Brain homing pepti	
42	4	19.0	8	19	AAV90957	Collagen derived p	
43	4	19.0	8	19	AAV90958	Collagen derived p	
44	4	19.0	8	19	AAV20879	Human presenilin I	
45	4	19.0	8	20	AAV40757	S5 derivative #10,	

# ALIGNMENTS

RESULT 1

AAW13415

ID AAW13415 standard; Peptide; 21 AA.

XX

AC AAW13415;

XX

DT 15-JAN-1998 (first entry)

XX

DE Brain homing peptide.

XX

KW Brain homing peptide; in vivo panning; screening; phage display;

KW drug delivery.

XX

OS Synthetic.

XX

PN WO9710507-A1.

XX

PD 20-MAR-1997.

XX

PF 10-SEP-1996; 96WO-US14600.

XX

PR 11-SEP-1995; 95US-0526710.

PR 11-SEP-1995; 95US-0526708.

XX

PA (LJOL-) LA JOLLA CANCER RES FOUND.

XX

PI Pasqualini R, Ruoslahti E;

XX

DR WPI; 1997-202359/18.

XX

PT Obtaining compound that homes to selected organ or tissue - by in

PT vivo panning method, specifically to identify brain, kidney,

PT angiogenic vasculature or tumour tissue homing peptide(s)

XX

PS Claim 13; Page 67; 75pp; English.

XX

CC This synthetic peptide is a claimed example of a brain-homing  
 CC peptide that was identified using a novel method for obtaining  
 CC molecules that home to a selected organ or tissue. This in vivo  
 CC panning method typically involves administering a phage display  
 CC library to a subject, and identifying expressed peptides which  
 CC home to the desired organ or tissue, e.g. brain, kidney, angiogenic  
 CC vascular tissue or tumour tissue. The isolated peptides (see  
 CC AAW13412-52, AAW11181-86) can be used to target e.g. drugs, toxins or  
 CC labels to the selected organ/tissue (claimed) or to identify and/or  
 CC isolate target molecules (claimed). The peptides can be directly  
 CC identified in vivo, as compared to prior art in vitro screening  
 CC methods, which require further examination to see if they maintain  
 CC specificity in vivo.

XX

SQ Sequence 21 AA;

Query Match 100.0%; Score 21; DB 18; Length 21;

Best Local Similarity 100.0%; Pred. No. 1.1e-14;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WRCVLREGPAGGCAWFNRHRL 21

Db

|||||  
1 WRCVLREGPAGGCAWFNRHRL 21

RESULT 2

AAB12002

ID AAB12002 standard; peptide; 21 AA.

XX

AC AAB12002;

XX

DT 17-OCT-2000 (first entry)

XX

DE Brain homing peptide # 16.

XX

KW Brain; homing peptide; organ targeting; tissue targeting; mouse; cyclic.

XX

OS Mus sp.

XX

PN US6068829-A.

XX

PD 30-MAY-2000.

XX

PF 23-JUN-1997; 97US-0862855.

XX

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Pasqualini R, Ruoslahti E;

XX

DR WPI; 2000-410850/35.

XX

PT Identifying and recovering organ homing molecules or peptides by in  
PT vivo panning comprises administering a library of diverse peptides  
PT linked to a tag which facilitates recovery of these peptides -

XX

PS Example 2; Column 17; 20pp; English.

XX

CC The present sequence is a mouse brain homing peptide. This sequence was  
CC identified by using in vivo panning to screen a library of potential  
CC organ homing molecules. The present sequence can be used to direct a  
CC moiety to a the brain tissue, by linking the moiety to the present  
CC sequence. Examples of potential moieties are drugs, toxins or a  
CC detectable label. The present sequence contains a VRL amino acid motif.

XX

SQ Sequence 21 AA;

Query Match 100.0%; Score 21; DB 21; Length 21;

Best Local Similarity 100.0%; Pred. No. 1.1e-14;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WRCVLREGPAGGCAWFNRHRL 21

|||||

Db 1 WRCVLREGPAGGCAWFNRHRL 21

RESULT 3

AAE11808

ID AAE11808 standard; peptide; 21 AA.

XX

AC AAE11808;

XX

DT 18-DEC-2001 (first entry)

XX

DE Phage peptide #16 targetted to brain.

XX

KW Enriched library fraction; brain; kidney; tumour; panning; diagnostic;  
KW molecular medicine; drug delivery; peptidomimetic; pharmaceutical.

XX

OS Bacteriophage.

XX

FH Key Location/Qualifiers

FT Domain 4..6

FT /label= VLR\_motif

XX

PN US6296832-B1.

XX

PD 02-OCT-2001.

XX

PF 08-JAN-1999; 99US-0226985.

XX

PR 23-JUN-1997; 97US-0862855.

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Ruoslahti E, Pasqualini R;

XX

DR WPI; 2001-610691/70.

XX

PT Enriched library fraction comprising molecules recovered by in vivo  
PT panning that selectively home to a selected organ or tissue useful for  
PT treating disease or in diagnostic methods -

XX

PS Example 2; Column 17; 21pp; English.

XX

CC The invention relates to an enriched library fraction containing  
CC molecules that selectively home to a selected organ or tissue such as  
CC brain, kidney or tumour recovered by in vivo panning. The invention  
CC generally relates to the field of molecular medicine, drug delivery and  
CC to a method of invivo panning for identifying a molecule that homes to a  
CC specific organ. The molecules, e.g., peptides, peptidomimetics, proteins  
CC and fragments of proteins contained in an enriched library fraction may  
CC be administered to a subject as part of a pharmaceutical composition to  
CC treat disease or in diagnostic methods. The present sequence is a  
CC peptide from bacteriophage targetted to brain.

XX

SQ Sequence 21 AA;

Query Match 100.0%; Score 21; DB 22; Length 21;

Best Local Similarity 100.0%; Pred. No. 1.1e-14;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WRCVLREGPAGGCAWFNRHRL 21  
 |||||  
 Db 1 WRCVLREGPAGGCAWFNRHRL 21

RESULT 4

AAU10719

ID AAU10719 standard; peptide; 21 AA.

XX

AC AAU10719;

XX

DT 12-MAR-2002 (first entry)

XX

DE Brain homing peptide #16 useful for delivery of target molecules.

XX

KW Organ targeting; tissue targeting; cancer; tumour homing molecule;

KW delivery of target molecule; brain homing peptide.

XX

OS Synthetic.

XX

PN US6306365-B1.

XX

PD 23-OCT-2001.

XX

PF 08-JAN-1999; 99US-0227906.

XX

PR 23-JUN-1997; 97US-0862855.

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX

PA (BURN-) BURNHAM INST.

XX

PI Ruoslahti E, Pasqualini R;

XX

DR WPI; 2002-040196/05.

XX

PT Recovering molecules that home to an organ or tissue, useful for  
 PT identifying molecules that home to a specific organ or tissue, e.g.  
 PT identifying a tumour homing molecule to identify the presence of cancer,  
 PT by in vivo panning of a library -

XX

PS Example 2; Column 17; 21pp; English.

XX

CC The present invention relates to a method of recovering molecules that  
 CC home to a selected organ or tissue. The method comprises administering  
 CC to the subject the library of diverse molecules, collecting a sample of  
 CC the selected organ or tissue (e.g. brain or kidney), and recovering from  
 CC the sample several molecules that home to the selected organ or tissue.  
 CC The method is useful for identifying molecules, particularly useful for  
 CC screening large number of molecules (e.g. peptides), that home to a  
 CC specific organ. The identified molecule is useful for e.g. raising an  
 CC antibody specific for a target molecule, targeting a desired moiety  
 CC (e.g. drug, toxin or detectable label) to the selected organ.  
 CC Specifically, the method is useful for identifying the presence of cancer  
 CC in a subject by linking an appropriate moiety to a tumour homing  
 CC molecule. The present method provides a direct means for identifying

CC molecules that specifically home to a selected organ and, therefore  
CC provides a significant advantage over previous methods, which require  
CC that a molecule identified using an in vitro screening method  
CC subsequently be examined to determine if it maintains its specificity in  
CC vivo. AAU10704-AAU10723 represent brain homing peptides described in  
CC the present invention.

XX

SQ Sequence 21 AA;

Query Match 100.0%; Score 21; DB 23; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.1e-14;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WRCVLREGPAGGCAWFNRHRL 21  
|||||  
Db 1 WRCVLREGPAGGCAWFNRHRL 21

# RESULT 5

ABU59531

ID ABU59531 standard; Peptide; 21 AA.

XX

AC ABU59531;

XX

DT 22-APR-2003 (first entry)

XX

DE Brain receptor targeting peptide #3.

XX

KW Targeting ligand; bioactive agent; polymer matrix; cancer; cytostatic;  
KW cathepsin-D substrate; peptides; neuroreceptor; adrenal receptor;  
KW fibronectin; vitronectin; integrin; RGD motif; angiogenic endothelium;  
KW tumour; cationic cancer-targeting peptide.

XX

OS Synthetic.

XX

PN US2002041898-A1.

XX

PD 11-APR-2002.

XX

PF 25-JUL-2001; 2001US-0912609.

XX

PR 05-JAN-2000; 2000US-0478124.

PR 31-OCT-2000; 2000US-0703474.

XX

PA (UNGE/) UNGER E C.

PA (MATS/) MATSUNAGA T O.

PA (RAMA/) RAMASWAMI V.

PA (ROMA/) ROMANOWSKI M J.

XX

PI Unger EC, Matsunaga TO, Ramaswami V, Romanowski MJ;

XX

DR WPI; 2003-208921/20.

XX

PT Targeted delivery system comprising a bioactive agent homogeneously  
PT dispersed in a targeted matrix is especially useful in cancer therapy  
PT -

XX



PS Claim 23; Page 37; 46pp; English.

XX

CC The invention relates to a composition comprising a bioactive agent  
CC homogeneously dispersed in a targeted matrix (polymer and targeting  
CC ligand). Also included are a targeted matrix for use as a delivery  
CC vehicle comprising a polymer associated with a targeting ligand,  
CC enhancing the bioavailability of an agent comprising administration  
CC of the composition and treating cancer comprising administration of the  
CC novel composition. The method is useful for targeted delivery of a drug,  
CC especially in cancer therapy. The targeting ligand may be a peptide.  
CC Examples of targeting peptides are disclosed including cathepsin-D  
CC substrate peptides, peptides targeting receptors in the brain and  
CC kidney, peptides recognising fibronectin- and vitronectin-binding  
CC integrins, peptides targeting the RGD (Arg-Gly-Asp)-motif in, e.g.,  
CC antibodies, peptides targeting the angiogenic endothelium of solid  
CC tumours, tissue specific peptides (e.g. of lung, skin, pancreas,  
CC intestine, uterus, adrenal gland and retina), and cationic cancer-  
CC targeting peptides. The present sequence is a peptide targeting  
CC ligand disclosed in the invention.

XX

SQ Sequence 21 AA;

Query Match 100.0%; Score 21; DB 24; Length 21;

Best Local Similarity 100.0%; Pred. No. 1.1e-14;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WRCVLREGPAGGCAWFNRHRL 21

||||||||||||||||||||

Db 1 WRCVLREGPAGGCAWFNRHRL 21

# RESULT 6

AAW13435

ID AAW13435 standard; Peptide; 9 AA.

XX

AC AAW13435;

XX

DT 15-JAN-1998 (first entry)

XX

DE Brain homing peptide.

XX

KW Brain homing peptide; in vivo panning; screening; phage display;  
KW drug delivery.

XX

OS Synthetic.

XX

PN WO9710507-A1.

XX

PD 20-MAR-1997.

XX

PF 10-SEP-1996; 96WO-US14600.

XX

PR 11-SEP-1995; 95US-0526710.

PR 11-SEP-1995; 95US-0526708.

XX

PA (LJOL-) LA JOLLA CANCER RES FOUND.

XX

PI Pasqualini R, Ruoslahti E;  
 XX  
 DR WPI; 1997-202359/18.  
 XX  
 PT Obtaining compound that homes to selected organ or tissue - by in  
 PT vivo panning method, specifically to identify brain, kidney,  
 PT angiogenic vasculature or tumour tissue homing peptide(s)  
 XX  
 PS Disclosure; Page 45; 75pp; English.  
 XX  
 CC This synthetic peptide is an example of a brain-homing peptide  
 CC that was identified using a claimed method for obtaining  
 CC molecules that home to a selected organ or tissue. This in vivo  
 CC panning method typically involves administering a phage display  
 CC library to a subject, and identifying expressed peptides which  
 CC home to the desired organ or tissue, e.g. brain, kidney, angiogenic  
 CC vascular tissue or tumour tissue. The isolated peptides (see  
 CC AAW13412-52, AAW11181-86) can be used to target e.g. drugs, toxins or  
 CC labels to the selected organ/tissue (claimed) or to identify and/or  
 CC isolate target molecules (claimed). The peptides can be directly  
 CC identified in vivo, as compared to prior art in vitro screening  
 CC methods, which require further examination to see if they maintain  
 CC specificity in vivo.  
 XX  
 SQ Sequence 9 AA;  
  
 Query Match 42.9%; Score 9; DB 18; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VLREGPAGG 12  
 |||||  
 Db 1 VLREGPAGG 9

# RESULT 7

AAB12006

ID AAB12006 standard; peptide; 9 AA.

XX

AC AAB12006;

XX

DT 17-OCT-2000 (first entry)

XX

DE Brain homing peptide # 20.

XX

KW Brain; homing peptide; organ targeting; tissue targeting; mouse; cyclic.

XX

OS Mus sp.

XX

PN US6068829-A.

XX

PD 30-MAY-2000.

XX

PF 23-JUN-1997; 97US-0862855.

XX

PR 11-SEP-1995; 95US-0526710.

PR 10-MAR-1997; 97US-0813273.

XX  
PA (BURN-) BURNHAM INST.  
XX  
PI Pasqualini R, Ruoslahti E;  
XX  
DR WPI; 2000-410850/35.  
XX  
PT Identifying and recovering organ homing molecules or peptides by in  
PT vivo panning comprises administering a library of diverse peptides  
PT linked to a tag which facilitates recovery of these peptides -  
XX  
PS Disclosure; Column 14; 20pp; English.  
XX  
CC The present sequence is a mouse brain homing peptide. This sequence was  
CC identified by using in vivo panning to screen a library of potential  
CC organ homing molecules. The present sequence can be used to direct a  
CC moiety to a the brain tissue, by linking the moiety to the present  
CC sequence. Examples of potential moieties are drugs, toxins or a  
CC detectable label.  
XX  
SQ Sequence 9 AA;

Query Match 42.9%; Score 9; DB 21; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VLREGPAGG 12  
| | | | | | | |  
Db 1 VLREGPAGG 9

RESULT 8  
AAE11812

ID AAE11812 standard; peptide; 9 AA.  
XX  
AC AAE11812;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Phage peptide #20 targetted to brain.  
XX  
KW Enriched library fraction; brain; kidney; tumour; panning; diagnostic;  
KW molecular medicine; drug delivery; peptidomimetic; pharmaceutical.  
XX  
OS Bacteriophage.  
XX  
PN US6296832-B1.  
XX  
PD 02-OCT-2001.  
XX  
PF 08-JAN-1999; 99US-0226985.  
XX  
PR 23-JUN-1997; 97US-0862855.  
PR 11-SEP-1995; 95US-0526710.  
PR 10-MAR-1997; 97US-0813273.  
XX  
PA (BURN-) BURNHAM INST.

XX  
 PI Ruoslahti E, Pasqualini R;  
 XX  
 DR WPI; 2001-610691/70.  
 XX  
 PT Enriched library fraction comprising molecules recovered by in vivo  
 PT panning that selectively home to a selected organ or tissue useful for  
 PT treating disease or in diagnostic methods -  
 XX  
 PS Disclosure; Column 14; 21pp; English.  
 XX  
 CC The invention relates to an enriched library fraction containing  
 CC molecules that selectively home to a selected organ or tissue such as  
 CC brain, kidney or tumour recovered by in vivo panning. The invention  
 CC generally relates to the field of molecular medicine, drug delivery and  
 CC to a method of invivo panning for identifying a molecule that homes to a  
 CC specific organ. The molecules, e.g., peptides, peptidomimetics, proteins  
 CC and fragments of proteins contained in an enriched library fraction may  
 CC be administered to a subject as part of a pharmaceutical composition to  
 CC treat disease or in diagnostic methods. The present sequence is a  
 CC peptide from bacteriophage targetted to brain.  
 XX  
 SQ Sequence 9 AA;

Query Match 42.9%; Score 9; DB 22; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VLREGPAGG 12  
 |||||  
 Db 1 VLREGPAGG 9

RESULT 9  
 AAU10723  
 ID AAU10723 standard; peptide; 9 AA.  
 XX  
 AC AAU10723;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 DE Brain homing peptide #20 useful for delivery of target molecules.  
 XX  
 KW Organ targeting; tissue targeting; cancer; tumour homing molecule;  
 KW delivery of target molecule; brain homing peptide.  
 XX  
 OS Synthetic.  
 XX  
 PN US6306365-B1.  
 XX  
 PD 23-OCT-2001.  
 XX  
 PF 08-JAN-1999; 99US-0227906.  
 XX  
 PR 23-JUN-1997; 97US-0862855.  
 PR 11-SEP-1995; 95US-0526710.  
 PR 10-MAR-1997; 97US-0813273.

XX  
 PA (BURN-) BURNHAM INST.  
 XX  
 PI Ruoslahti E, Pasqualini R;  
 XX  
 DR WPI; 2002-040196/05.  
 XX  
 PT Recovering molecules that home to an organ or tissue, useful for  
 PT identifying molecules that home to a specific organ or tissue, e.g.  
 PT identifying a tumour homing molecule to identify the presence of cancer,  
 PT by in vivo panning of a library -  
 XX  
 PS Disclosure; Column 14; 2lpp; English.  
 XX  
 CC The present invention relates to a method of recovering molecules that  
 CC home to a selected organ or tissue. The method comprises administering  
 CC to the subject the library of diverse molecules, collecting a sample of  
 CC the selected organ or tissue (e.g. brain or kidney), and recovering from  
 CC the sample several molecules that home to the selected organ or tissue.  
 CC The method is useful for identifying molecules, particularly useful for  
 CC screening large number of molecules (e.g. peptides), that home to a  
 CC specific organ. The identified molecule is useful for e.g. raising an  
 CC antibody specific for a target molecule, targeting a desired moiety  
 CC (e.g. drug, toxin or detectable label) to the selected organ.  
 CC Specifically, the method is useful for identifying the presence of cancer  
 CC in a subject by linking an appropriate moiety to a tumour homing  
 CC molecule. The present method provides a direct means for identifying  
 CC molecules that specifically home to a selected organ and, therefore  
 CC provides a significant advantage over previous methods, which require  
 CC that a molecule identified using an in vitro screening method  
 CC subsequently be examined to determine if it maintains its specificity in  
 CC vivo. AAU10704-AAU10723 represent brain homing peptides described in  
 CC the present invention.  
 XX  
 SQ Sequence 9 AA;

Query Match 42.9%; Score 9; DB 23; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 VLREGPAGG 12  
 |||||  
 Db 1 VLREGPAGG 9

# RESULT 10

AA37974

ID AA37974 standard; Peptide; 9 AA.

XX

AC AA37974;

XX

DT 29-SEP-1999 (first entry)

XX

DE Human cERB2 oncogene-derived HLA-binding peptide.

XX

KW Immunogen; HLA; human leukocyte antigen; binding motif; antiviral;

KW MHC; major histocompatibility complex; viral infection; anticancer;

KW prostate cancer; lymphoma; hepatitis; AIDS; diagnostic; diagnosis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9403205-A1.  
 XX  
 PD 17-FEB-1994.  
 XX  
 PF 06-AUG-1993; 93WO-US07421.  
 XX  
 PR 05-MAR-1993; 93US-0027746.  
 PR 07-AUG-1992; 92US-0926666.  
 XX  
 PA (CYTE-) CYTEL CORP.  
 XX  
 PI Celis E, Grey HM, Kubo RT, Sette A;  
 XX  
 DR WPI; 1994-065403/08.  
 XX  
 PT Peptide which specifically binds selected MHC allele - used to  
 PT induce an immune response for treatment or prevention of viral  
 PT infection or cancer, or for diagnosis  
 XX  
 PS Disclosure; Page 103; 150pp; English.  
 XX  
 CC The sequence is a specific example of a group of new immunogenic  
 CC peptides having an HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1 binding  
 CC motif. For example, the peptides having an HLA-A3.2 binding motif  
 CC each have 9-10 residues and contain, from the N-terminus to the  
 CC C-terminus, (a) a first conserved residue selected from L, M, I,  
 CC V, S, A, T, F, C, G, D and E and (b) a second conserved residue of  
 CC K, R, Y, H or F, where the first and second conserved residues are  
 CC separated by 6-7 residues. The peptides are capable of binding  
 CC selected MHC molecules and inducing an immune response. They can be  
 CC used to treat and/or prevent viral infection and cancer, e.g. prostate  
 CC cancer, lymphoma, hepatitis or AIDS. They can also be used to produce  
 CC antibodies for use as diagnostic or therapeutic agents. The peptides  
 CC can also be used as diagnostic agents.  
 XX  
 SQ Sequence 9 AA;

Query Match 23.8%; Score 5; DB 15; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGGCA 14  
 |||||  
 Db 4 AGGCA 8

RESULT 11  
 AAY45544  
 ID AAY45544 standard; Peptide; 9 AA.  
 XX  
 AC AAY45544;  
 XX  
 DT 01-DEC-1999 (first entry)

XX  
 DE Immunogenic peptide having a human leukocyte antigen binding motif #155.  
 XX  
 KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;  
 KW immune response; T cell activation; major histocompatibility complex;  
 KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;  
 KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;  
 KW vaccine; immunisation.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9945954-A1.  
 XX  
 PD 16-SEP-1999.  
 XX  
 PF 13-MAR-1998; 98WO-US05039.  
 XX  
 PR 13-MAR-1998; 98WO-US05039.  
 XX  
 PA (EPIM-) EPIMMUNE INC.  
 XX  
 PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;  
 XX  
 DR WPI; 1999-551214/46.  
 XX  
 PT New immunogenic peptides with HLA binding motif, useful in treatment  
 PT and diagnosis of cancers and viral diseases -  
 XX  
 PS Claim 1; Page 33; 150pp; English.  
 XX  
 CC AAY45390 to AAY48214 represent specifically claimed immunogenic peptides  
 CC having a human major histocompatibility complex (MHC) Class I (also  
 CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic  
 CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes  
 CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell  
 CC response against the antigen from which the peptide is derived.  
 CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are  
 CC normally induced by an antigen in the form of a peptide fragment bound  
 CC to a HLA molecule, rather than the intact foreign antigen itself, and  
 CC are particularly important in tumour rejection and in fighting viral  
 CC infections. The peptides are therefore useful therapeutically to treat  
 CC or prevent viral infections and cancers in mammals (especially humans)  
 CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.  
 CC They can be administered as vaccines to elicit an immune response in  
 CC individuals susceptible or otherwise at risk of viral infection or  
 CC cancer, or used to treat chronic or acute conditions. They are also  
 CC useful diagnostically, and can be used to induce a cytotoxic T cell  
 CC response, by contacting a cytotoxic T cell with the peptide e.g. to  
 CC produce CTLs ex vivo for infusion back into a patient. The  
 CC polynucleotides encoding the immunogenic peptides are also useful  
 CC therapeutically and for immunisation as above.  
 XX  
 SQ Sequence 9 AA;

Query Match 23.8%; Score 5; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGGCA 14  
 |||||  
 Db 4 AGGCA 8

RESULT 12

AAG88625

ID AAG88625 standard; Peptide; 9 AA.

XX

AC AAG88625;

XX

DT 11-SEP-2001 (first entry)

XX

DE HER2/NEU DR supermotif binding peptide core sequence #169.

XX

KW Human; HER2/neu; epitope; human leukocyte antigen; HLA; T cell;  
 KW immune response; vaccine; tumour; cancer; cytostatic; immunostimulant;  
 KW tumour-associated antigen; T lymphocyte; cytotoxic T lymphocyte; CTL.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN WO200141787-A1.

XX

PD 14-JUN-2001.

XX

PF 11-DEC-2000; 2000WO-US33591.

XX

PR 10-DEC-1999; 99US-0458299.

XX

PA (EPIM-) EPIMMUNE INC.

XX

PI Fikes J, Sette A, Sidney J, Southwood S, Chesnut R, Celis E;  
 PI Keogh E;

XX

DR WPI; 2001-374995/39.

XX

PT An isolated prepared HER2/neu epitope useful in a vaccine for inducing  
 PT cellular immune responses for the prevention and treatment of cancer -

XX

PS Disclosure; Page 172; 199pp; English.

XX

CC The present invention describes isolated prepared HER2/neu epitopes (I).  
 CC Also described are: (1) a clonal cytotoxic T lymphocyte (CTL) that is  
 CC culture in vitro and binds to a complex of an epitope (I), bound to a  
 CC human leukocyte antigen (HLA) molecule; (2) a peptide (II) comprising (I)  
 CC and a second epitope and the peptide is less than 50 contiguous amino  
 CC acids that have 100% identity with a native peptide sequence of HER2/neu;  
 CC (3) a vaccine composition (III) comprising (II) and a pharmaceutical  
 CC excipient; (4) an isolated nucleic acid encoding a peptide comprising  
 CC (I); and (5) an isolated nucleic acid encoding (II). (I) has cytostatic  
 CC and immunostimulant activities, and can be used in vaccines. (I), (II)  
 CC and (III) are useful for inducing cellular immune responses for the  
 CC prevention and treatment of cancer. (I) and (II) are useful for  
 CC monitoring or evaluating an immune response to a tumour-associated



CC antigen when incubated with a T lymphocyte sample from a patient and  
 CC detecting the presence of bound T lymphocyte to (I) or (II). Epitope  
 CC based vaccines mean that immunosuppressive epitopes that may be present  
 CC in whole antigens may be avoided. Selected epitopes may be combined to  
 CC enhance immunogenicity. The possible pathological side effects caused by  
 CC infectious agents or whole protein antigen is eliminated. The vaccine  
 CC provides the ability to direct and focus an immune response to multiple  
 CC selected antigens from the same pathogen. Epitope-based anti-tumour  
 CC vaccines provides the opportunity to combine epitopes derived from  
 CC multiple tumour-associated molecules addressing the problem of tumour-  
 CC tumour variability and reducing the likelihood of tumour escape due to  
 CC antigen loss. AAG88266 to AAG89121 represent amino acid sequences used in  
 CC the exemplification of the present invention.

XX

SQ Sequence 9 AA;

Query Match 23.8%; Score 5; DB 22; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGGCA 14  
 |||||  
 Db 3 AGGCA 7

# RESULT 13

AAG88807

ID AAG88807 standard; Peptide; 9 AA.

XX

AC AAG88807;

XX

DT 11-SEP-2001 (first entry)

XX

DE HER2/neu epitope HLA-A3 supermotif-bearing peptide #11.

XX

KW Human; HER2/neu; epitope; human leukocyte antigen; HLA; T cell;  
 KW immune response; vaccine; tumour; cancer; cytostatic; immunostimulant;  
 KW tumour-associated antigen; T lymphocyte; cytotoxic T lymphocyte; CTL.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN WO200141787-A1.

XX

PD 14-JUN-2001.

XX

PF 11-DEC-2000; 2000WO-US33591.

XX

PR 10-DEC-1999; 99US-0458299.

XX

PA (EPIM-) EPIMMUNE INC.

XX

PI Fikes J, Sette A, Sidney J, Southwood S, Chesnut R, Celis E;  
 PI Keogh E;

XX

DR WPI; 2001-374995/39.

XX

PT An isolated prepared HER2/neu epitope useful in a vaccine for inducing  
PT cellular immune responses for the prevention and treatment of cancer -  
XX

PS Claim 1; Page 181; 199pp; English.  
XX

CC The present invention describes isolated prepared HER2/neu epitopes (I).  
CC Also described are: (1) a clonal cytotoxic T lymphocyte (CTL) that is  
CC culture in vitro and binds to a complex of an epitope (I), bound to a  
CC human leukocyte antigen (HLA) molecule; (2) a peptide (II) comprising (I)  
CC and a second epitope and the peptide is less than 50 contiguous amino  
CC acids that have 100% identity with a native peptide sequence of HER2/neu;  
CC (3) a vaccine composition (III) comprising (II) and a pharmaceutical  
CC excipient; (4) an isolated nucleic acid encoding a peptide comprising  
CC (I); and (5) an isolated nucleic acid encoding (II). (I) has cytostatic  
CC and immunostimulant activities, and can be used in vaccines. (I), (II)  
CC and (III) are useful for inducing cellular immune responses for the  
CC prevention and treatment of cancer. (I) and (II) are useful for  
CC monitoring or evaluating an immune response to a tumour-associated  
CC antigen when incubated with a T lymphocyte sample from a patient and  
CC detecting the presence of bound T lymphocyte to (I) or (II). Epitope  
CC based vaccines mean that immunosuppressive epitopes that may be present  
CC in whole antigens may be avoided. Selected epitopes may be combined to  
CC enhance immunogenicity. The possible pathological side effects caused by  
CC infectious agents or whole protein antigen is eliminated. The vaccine  
CC provides the ability to direct and focus an immune response to multiple  
CC selected antigens from the same pathogen. Epitope-based anti-tumour  
CC vaccines provides the opportunity to combine epitopes derived from  
CC multiple tumour-associated molecules addressing the problem of tumour-  
CC tumour variability and reducing the likelihood of tumour escape due to  
CC antigen loss. AAG88266 to AAG89121 represent amino acid sequences used in  
CC the exemplification of the present invention.  
XX

SQ Sequence 9 AA;

Query Match 23.8%; Score 5; DB 22; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGGCA 14  
|||  
Db 4 AGGCA 8

#### RESULT 14

AAY37995

ID AAY37995 standard; Peptide; 10 AA.

XX

AC AAY37995;

XX

DT 29-SEP-1999 (first entry)

XX

DE Human cERB2 oncogene-derived HLA-binding peptide.

XX

KW Immunogen; HLA; human leukocyte antigen; binding motif; antiviral;  
KW MHC; major histocompatibility complex; viral infection; anticancer;  
KW prostate cancer; lymphoma; hepatitis; AIDS; diagnostic; diagnosis.

XX

OS Homo sapiens.  
 XX  
 PN WO9403205-A1.  
 XX  
 PD 17-FEB-1994.  
 XX  
 PF 06-AUG-1993; 93WO-US07421.  
 XX  
 PR 05-MAR-1993; 93US-0027746.  
 PR 07-AUG-1992; 92US-0926666.  
 XX  
 PA (CYTE-) CYTEL CORP.  
 XX  
 PI Celis E, Grey HM, Kubo RT, Sette A;  
 XX  
 DR WPI; 1994-065403/08.  
 XX  
 PT Peptide which specifically binds selected MHC allele - used to  
 PT induce an immune response for treatment or prevention of viral  
 PT infection or cancer, or for diagnosis  
 XX  
 PS Disclosure; Page 103; 150pp; English.  
 XX  
 CC The sequence is a specific example of a group of new immunogenic  
 CC peptides having an HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1 binding  
 CC motif. For example, the peptides having an HLA-A3.2 binding motif  
 CC each have 9-10 residues and contain, from the N-terminus to the  
 CC C-terminus, (a) a first conserved residue selected from L, M, I,  
 CC V, S, A, T, F, C, G, D and E and (b) a second conserved residue of  
 CC K, R, Y, H or F, where the first and second conserved residues are  
 CC separated by 6-7 residues. The peptides are capable of binding  
 CC selected MHC molecules and inducing an immune response. They can be  
 CC used to treat and/or prevent viral infection and cancer, e.g. prostate  
 CC cancer, lymphoma, hepatitis or AIDS. They can also be used to produce  
 CC antibodies for use as diagnostic or therapeutic agents. The peptides  
 CC can also be used as diagnostic agents.  
 XX  
 SQ Sequence 10 AA;  
  
 Query Match 23.8%; Score 5; DB 15; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 10 AGGCA 14  
 |||||  
 Db 5 AGGCA 9

RESULT 15  
 AAY45565  
 ID AAY45565 standard; Peptide; 10 AA.  
 XX  
 AC AAY45565;  
 XX  
 DT 01-DEC-1999 (first entry)  
 XX  
 DE Immunogenic peptide having a human leukocyte antigen binding motif #176.

XX  
 KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;  
 KW immune response; T cell activation; major histocompatibility complex;  
 KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;  
 KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;  
 KW vaccine; immunisation.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN WO9945954-A1.  
 XX  
 PD 16-SEP-1999.  
 XX  
 PF 13-MAR-1998; 98WO-US05039.  
 XX  
 PR 13-MAR-1998; 98WO-US05039.  
 XX  
 PA (EPIM-) EPIMMUNE INC.  
 XX  
 PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;  
 XX  
 DR WPI; 1999-551214/46.  
 XX  
 PT New immunogenic peptides with HLA binding motif, useful in treatment  
 PT and diagnosis of cancers and viral diseases -  
 XX  
 PS Claim 1; Page 34; 150pp; English.  
 XX  
 CC AAY45390 to AAY48214 represent specifically claimed immunogenic peptides  
 CC having a human major histocompatibility complex (MHC) Class I (also  
 CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic  
 CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes  
 CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell  
 CC response against the antigen from which the peptide is derived.  
 CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are  
 CC normally induced by an antigen in the form of a peptide fragment bound  
 CC to a HLA molecule, rather than the intact foreign antigen itself, and  
 CC are particularly important in tumour rejection and in fighting viral  
 CC infections. The peptides are therefore useful therapeutically to treat  
 CC or prevent viral infections and cancers in mammals (especially humans)  
 CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.  
 CC They can be administered as vaccines to elicit an immune response in  
 CC individuals susceptible or otherwise at risk of viral infection or  
 CC cancer, or used to treat chronic or acute conditions. They are also  
 CC useful diagnostically, and can be used to induce a cytotoxic T cell  
 CC response, by contacting a cytotoxic T cell with the peptide e.g. to  
 CC produce CTLs ex vivo for infusion back into a patient. The  
 CC polynucleotides encoding the immunogenic peptides are also useful  
 CC therapeutically and for immunisation as above.  
 XX  
 SQ Sequence 10 AA;

Query Match 23.8%; Score 5; DB 20; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      10 AGGCA 14
          |||||
Db       5 AGGCA 9

```

Search completed: November 13, 2003, 10:32:57  
Job time : 68.375 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 10:38:27 ; Search time 41.125 Seconds  
(without alignments)  
93.222 Million cell updates/sec

Title: US-09-228-866-16  
Perfect score: 21  
Sequence: 1 WRCVLREGPAGGCAWFNRHRL 21

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 666188 seqs, 182559486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 124183

```
Minimum DB seq length: 7
Maximum DB seq length: 21
```

Post-processing: Listing first 45 summaries

```
Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	9	42.9	9	12	US-10-306-878-12		Sequence 12, Appl
2	5	23.8	10	9	US-09-834-765-412		Sequence 412, App
3	5	23.8	10	9	US-09-834-765-573		Sequence 573, App
4	5	23.8	10	9	US-09-834-765-690		Sequence 690, App
5	5	23.8	10	11	US-09-572-404B-3780		Sequence 3780, Ap
6	5	23.8	12	12	US-10-083-768-85		Sequence 85, Appl
7	5	23.8	18	10	US-09-784-199-10		Sequence 10, Appl
8	4	19.0	7	11	US-09-792-686A-30		Sequence 30, Appl
9	4	19.0	7	12	US-10-022-066-638		Sequence 638, App
10	4	19.0	7	12	US-10-006-760-34		Sequence 34, Appl
11	4	19.0	8	10	US-09-756-283A-48		Sequence 48, Appl
12	4	19.0	8	10	US-09-756-283A-50		Sequence 50, Appl
13	4	19.0	8	10	US-09-756-283A-52		Sequence 52, Appl
14	4	19.0	8	10	US-09-756-283A-53		Sequence 53, Appl
15	4	19.0	8	10	US-09-848-967-29		Sequence 29, Appl
16	4	19.0	8	12	US-09-800-187-44		Sequence 44, Appl
17	4	19.0	8	12	US-10-280-066-204		Sequence 204, App
18	4	19.0	9	9	US-09-760-599-42		Sequence 42, Appl
19	4	19.0	9	10	US-09-017-743C-109		Sequence 109, App
20	4	19.0	9	11	US-09-938-864-92		Sequence 92, Appl
21	4	19.0	9	11	US-09-938-864-134		Sequence 134, App
22	4	19.0	9	11	US-09-938-864-215		Sequence 215, App
23	4	19.0	9	11	US-09-791-477-92		Sequence 92, Appl
24	4	19.0	9	11	US-09-791-477-134		Sequence 134, App
25	4	19.0	9	11	US-09-791-477-215		Sequence 215, App
26	4	19.0	9	11	US-09-785-019-92		Sequence 92, Appl
27	4	19.0	9	11	US-09-785-019-134		Sequence 134, App
28	4	19.0	9	11	US-09-785-019-215		Sequence 215, App
29	4	19.0	9	12	US-10-141-645-93		Sequence 93, Appl
30	4	19.0	9	12	US-10-141-645-105		Sequence 105, App
31	4	19.0	9	12	US-10-141-645-112		Sequence 112, App
32	4	19.0	9	12	US-10-141-645-113		Sequence 113, App
33	4	19.0	9	12	US-10-141-645-118		Sequence 118, App
34	4	19.0	9	12	US-10-141-645-119		Sequence 119, App
35	4	19.0	9	12	US-10-062-109A-286		Sequence 286, App
36	4	19.0	9	12	US-10-062-109A-356		Sequence 356, App
37	4	19.0	9	12	US-10-062-109A-656		Sequence 656, App
38	4	19.0	9	12	US-09-942-052-16		Sequence 16, Appl
39	4	19.0	9	12	US-09-942-052-44		Sequence 44, Appl
40	4	19.0	9	12	US-09-942-052-247		Sequence 247, App
41	4	19.0	9	12	US-09-942-052-411		Sequence 411, App
42	4	19.0	9	12	US-09-942-052-516		Sequence 516, App
43	4	19.0	9	12	US-09-942-052-546		Sequence 546, App
44	4	19.0	9	12	US-09-942-052-607		Sequence 607, App
45	4	19.0	9	12	US-10-239-313A-689		Sequence 689, App

#### ALIGNMENTS

RESULT 1

US-10-306-878-12

; Sequence 12, Application US/10306878  
 ; Publication No. US20030175819A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reed, John C.  
 ; APPLICANT: Guo, Bin  
 ; TITLE OF INVENTION: Methods for Identifying Modulators of  
 ; TITLE OF INVENTION: Apoptosis  
 ; FILE REFERENCE: P-LJ 5535  
 ; CURRENT APPLICATION NUMBER: US/10/306,878  
 ; CURRENT FILING DATE: 2002-11-27  
 ; PRIOR APPLICATION NUMBER: US 60/334,149  
 ; PRIOR FILING DATE: 2001-11-28  
 ; NUMBER OF SEQ ID NOS: 28  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 12  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Synthetic construct

US-10-306-878-12

Query Match 42.9%; Score 9; DB 12; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 6e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VLREGPAGG 12  
 |||||  
 Db 1 VLREGPAGG 9

RESULT 2

US-09-834-765-412

; Sequence 412, Application US/09834765  
 ; Patent No. US20020055478A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mary Faris  
 ; APPLICANT: Pia M. Challita-Eid  
 ; APPLICANT: Arthur B. Raitano  
 ; APPLICANT: Steve Chappell Mitchell  
 ; APPLICANT: Daniel E.H. Afar  
 ; APPLICANT: Aya Jakobovits  
 ; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
 ; TITLE OF INVENTION: AND DETECTION OF CANCER  
 ; FILE REFERENCE: 129.6USU1  
 ; CURRENT APPLICATION NUMBER: US/09/834,765  
 ; CURRENT FILING DATE: 2001-09-21  
 ; PRIOR APPLICATION NUMBER: 60/197,647  
 ; PRIOR FILING DATE: 2000-04-12  
 ; NUMBER OF SEQ ID NOS: 770  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 412  
 ; LENGTH: 10  
 ; TYPE: PRT

; ORGANISM: Homo sapiens  
US-09-834-765-412

Query Match 23.8%; Score 5; DB 9; Length 10;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 REGPA 10  
|||||  
Db 4 REGPA 8

#### RESULT 3

US-09-834-765-573  
; Sequence 573, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Faris  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar  
; APPLICANT: Aya Jakobovits  
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT  
; TITLE OF INVENTION: AND DETECTION OF CANCER  
; FILE REFERENCE: 129.6USU1  
; CURRENT APPLICATION NUMBER: US/09/834,765  
; CURRENT FILING DATE: 2001-09-21  
; PRIOR APPLICATION NUMBER: 60/197,647  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 770  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 573  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-834-765-573

Query Match 23.8%; Score 5; DB 9; Length 10;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 REGPA 10  
|||||  
Db 6 REGPA 10

#### RESULT 4

US-09-834-765-690  
; Sequence 690, Application US/09834765  
; Patent No. US20020055478A1  
; GENERAL INFORMATION:  
; APPLICANT: Mary Faris  
; APPLICANT: Pia M. Challita-Eid  
; APPLICANT: Arthur B. Raitano  
; APPLICANT: Steve Chappell Mitchell  
; APPLICANT: Daniel E.H. Afar



```
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT
; TITLE OF INVENTION: AND DETECTION OF CANCER
; FILE REFERENCE: 129.6USU1
; CURRENT APPLICATION NUMBER: US/09/834,765
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/197,647
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 770
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 690
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-765-690
```

```
Query Match          23.8%; Score 5; DB 9; Length 10;
Best Local Similarity 100.0%; Pred. No. 83;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      6 REGPA 10
        |||||
Db      6 REGPA 10
```

#### RESULT 5

```
US-09-572-404B-3780
; Sequence 3780, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3780
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in ERBB2 OR HER2 OR NGL OR NEU at 218-
227 and may
; OTHER INFORMATION: interact with Sequence 3781 in this patent.
US-09-572-404B-3780
```

```
Query Match          23.8%; Score 5; DB 11; Length 10;
Best Local Similarity 100.0%; Pred. No. 83;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      10 AGGCA 14
        |||||
Db      4 AGGCA 8
```

#### RESULT 6

US-10-083-768-85  
 ; Sequence 85, Application US/10083768  
 ; Publication No. US20030158116A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dower, William J.  
 ; Barrett, Ronald W.  
 ; Cwirla, Steven E.  
 ; Duffin, David J.  
 ; Gates, Christian  
 ; Haselden, Sherril S.  
 ; Mattheakis, Larry C.  
 ; Schatz, Peter J.  
 ; Wagstrom, Christopher R.  
 ; Wrighton, Nicholas C.  
 ; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
 ; THROMBOPOIETIN RECEPTOR  
 ; NUMBER OF SEQUENCES: 232  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Glaxo Wellcome  
 ; STREET: Five Moore Drive, P.O. Box 13398  
 ; CITY: Research Triangle Park  
 ; STATE: NC  
 ; COUNTRY: USA  
 ; ZIP: 27709  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/083,768  
 ; FILING DATE: 27-Feb-2002  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Hrubiec, Robert T.  
 ; REGISTRATION NUMBER: 36,392  
 ; REFERENCE/DOCKET NUMBER: PK3065USW  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 919-248-1000  
 ; INFORMATION FOR SEQ ID NO: 85:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 12 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: <Unknown>  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 85:  
 US-10-083-768-85

Query Match 23.8%; Score 5; DB 12; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 96;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CVLRE 7  
 |||||  
 Db 1 CVLRE 5

RESULT 7

US-09-784-199-10

```
; Sequence 10, Application US/09784199
; Patent No. US20020164738A1
; GENERAL INFORMATION:
; APPLICANT: MORISHIMA, No. US20020164738A1uhiro
; APPLICANT: SHIBATA, Takehiko
; TITLE OF INVENTION: CELL DEATH INHIBITORY PROTEIN
; FILE REFERENCE: 04853-0057-00000
; CURRENT APPLICATION NUMBER: US/09/784,199
; CURRENT FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: JP 2000-41927
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic peptide
US-09-784-199-10
```

```
Query Match          23.8%; Score 5; DB 10; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches      5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      4 VLREG 8
        |||||
Db      12 VLREG 16
```

RESULT 8

US-09-792-686A-30

```
; Sequence 30, Application US/09792686A
; Publication No. US20030068645A1
; GENERAL INFORMATION:
; APPLICANT: Goronzy, Jorg J.
; APPLICANT: Kopecky, Stephen L.
; APPLICANT: Weyand, Cornelia M.
; TITLE OF INVENTION: METHODS AND MATERIALS FOR EVALUATING
; TITLE OF INVENTION: CARDIOVASCULAR CONDITIONS
; FILE REFERENCE: 07039-250001
; CURRENT APPLICATION NUMBER: US/09/792,686A
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-792-686A-30
```

```
Query Match          19.0%; Score 4; DB 11; Length 7;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 9 PAGG 12  
 ||||  
Db 2 PAGG 5

RESULT 9

US-10-022-066-638

; Sequence 638, Application US/10022066  
; Publication No. US20030166057A1  
; GENERAL INFORMATION:  
; APPLICANT: HILDEBRAND, WILLIAM H.  
; APPLICANT: PRILLIMAN, KILEY RAE  
; TITLE OF INVENTION: METHOD AND APPARATUS FOR THE PRODUCTION OF ANTIGENS AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: 6680.034  
; CURRENT APPLICATION NUMBER: US/10/022,066  
; CURRENT FILING DATE: 2002-09-09  
; PRIOR APPLICATION NUMBER: 60/256,410  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: 60/256,409  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: 09/465,321  
; PRIOR FILING DATE: 1999-12-17  
; PRIOR APPLICATION NUMBER: 09/974,366  
; PRIOR FILING DATE: 2001-10-10  
; NUMBER OF SEQ ID NOS: 638  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 638  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-10-022-066-638

Query Match 19.0%; Score 4; DB 12; Length 7;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 PAGG 12  
 ||||  
Db 4 PAGG 7

RESULT 10

US-10-006-760-34

; Sequence 34, Application US/10006760  
; Publication No. US20030186385A1  
; GENERAL INFORMATION:  
; APPLICANT: Koide, Shohei  
; TITLE OF INVENTION: METHOD OF IDENTIFYING POLYPEPTIDE MONOBODIES WHICH BIND  
; TITLE OF INVENTION: TO TARGET PROTEINS AND USE THEREOF  
; FILE REFERENCE: 176/60901  
; CURRENT APPLICATION NUMBER: US/10/006,760  
; CURRENT FILING DATE: 2001-11-19  
; PRIOR APPLICATION NUMBER: 60/249,756

; PRIOR FILING DATE: 2000-11-17  
; NUMBER OF SEQ ID NOS: 73  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: AB loop  
; OTHER INFORMATION: sequence for polypeptide monobody in pYT45AB7N  
; OTHER INFORMATION: library  
US-10-006-760-34

Query Match 19.0%; Score 4; DB 12; Length 7;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VLRE 7  
|||  
Db 4 VLRE 7

RESULT 11

US-09-756-283A-48  
; Sequence 48, Application US/09756283A  
; Patent No. US20020151478A1  
; GENERAL INFORMATION:  
; APPLICANT: Chernajovsky, Yuti  
; APPLICANT: Dreja, Hanna Stina  
; APPLICANT: Adams, Gillian  
; TITLE OF INVENTION: Latent Fusion Protein  
; FILE REFERENCE: 0623.1000000  
; CURRENT APPLICATION NUMBER: US/09/756,283A  
; CURRENT FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 100  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 48  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-09-756-283A-48

Query Match 19.0%; Score 4; DB 10; Length 8;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GPAG 11  
|||  
Db 1 GPAG 4

RESULT 12

US-09-756-283A-50  
; Sequence 50, Application US/09756283A  
; Patent No. US20020151478A1  
; GENERAL INFORMATION:  
; APPLICANT: Chernajovsky, Yuti

```

; APPLICANT: Dreja, Hanna Stina
; APPLICANT: Adams, Gillian
; TITLE OF INVENTION: Latent Fusion Protein
; FILE REFERENCE: 0623.1000000
; CURRENT APPLICATION NUMBER: US/09/756,283A
; CURRENT FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-756-283A-50

```

```

Query Match          19.0%; Score 4; DB 10; Length 8;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      8 GPAG 11
        ||||
Db      1 GPAG 4

```

#### RESULT 13

```

US-09-756-283A-52
; Sequence 52, Application US/09756283A
; Patent No. US20020151478A1
; GENERAL INFORMATION:
; APPLICANT: Chernajovsky, Yuti
; APPLICANT: Dreja, Hanna Stina
; APPLICANT: Adams, Gillian
; TITLE OF INVENTION: Latent Fusion Protein
; FILE REFERENCE: 0623.1000000
; CURRENT APPLICATION NUMBER: US/09/756,283A
; CURRENT FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-756-283A-52

```

```

Query Match          19.0%; Score 4; DB 10; Length 8;
Best Local Similarity 100.0%; Pred. No. 6e+05;
Matches      4; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      8 GPAG 11
        ||||
Db      1 GPAG 4

```

#### RESULT 14

```

US-09-756-283A-53
; Sequence 53, Application US/09756283A
; Patent No. US20020151478A1
; GENERAL INFORMATION:

```

; APPLICANT: Chernajovsky, Yuti  
; APPLICANT: Dreja, Hanna Stina  
; APPLICANT: Adams, Gillian  
; TITLE OF INVENTION: Latent Fusion Protein  
; FILE REFERENCE: 0623.1000000  
; CURRENT APPLICATION NUMBER: US/09/756,283A  
; CURRENT FILING DATE: 2001-01-09  
; NUMBER OF SEQ ID NOS: 100  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 53  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-09-756-283A-53

Query Match 19.0%; Score 4; DB 10; Length 8;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GPAG 11  
|||  
Db 1 GPAG 4

RESULT 15

US-09-848-967-29

; Sequence 29, Application US/09848967  
; Publication No. US20020193295A1  
; GENERAL INFORMATION:  
; APPLICANT: CALENOFF, EMANUEL  
; APPLICANT: DITLOW, CHARLES C.  
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES AND USES THEREOF  
; FILE REFERENCE: 21417/92141  
; CURRENT APPLICATION NUMBER: US/09/848,967  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: 09/837,630  
; PRIOR FILING DATE: 2001-04-18  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 29  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-848-967-29

Query Match 19.0%; Score 4; DB 10; Length 8;  
Best Local Similarity 100.0%; Pred. No. 6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GPAG 11  
|||  
Db 3 GPAG 6

Search completed: November 13, 2003, 11:12:33  
Job time : 41.125 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 10:26:01 ; Search time 20.5625 Seconds  
(without alignments)  
98.215 Million cell updates/sec

Title: US-09-228-866-16  
Perfect score: 21  
Sequence: 1 WRCVLREGPAGGCAWFNRHRL 21

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3752

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : PIR\_76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	5	23.8	13	2	B36042	oxfX protein - Esc	
2	4	19.0	17	2	S16144	spoVA protein - Ba	
3	4	19.0	18	2	S46418	NTL1 protein - cur	
4	4	19.0	18	2	S19914	choline O-acetyltr	
5	4	19.0	19	2	B29501	fibrinopeptide A -	
6	4	19.0	21	2	C39543	collagen alpha 3(I	
7	3	14.3	7	2	PT0666	T-cell receptor be	
8	3	14.3	7	2	PT0683	T-cell receptor be	
9	3	14.3	8	2	PT0527	T-cell receptor be	
10	3	14.3	8	2	PT0509	T-cell receptor be	
11	3	14.3	8	4	I54017	granulocyte-colony	
12	3	14.3	9	2	S35538	ribosomal protein	
13	3	14.3	9	2	QDRB	delta sleep-induci	
14	3	14.3	9	2	A53797	3',5'-cyclic-GMP p	



15	3	14.3	9	2	S78420	ribosomal protein
16	3	14.3	10	2	S65388	cytochrome-c oxida
17	3	14.3	10	2	B33710	ornithine decarbox
18	3	14.3	10	2	H28027	protein P11 - curl
19	3	14.3	10	2	S77990	cytochrome-c oxida
20	3	14.3	10	2	S68638	acetylcholinestera
21	3	14.3	10	2	PT0212	T-cell receptor al
22	3	14.3	10	2	PC4374	telomeric and tetr
23	3	14.3	11	2	D57789	gallbladder stone
24	3	14.3	11	2	I41946	T-cell receptor ga
25	3	14.3	11	4	PC2124	aminotransferase c
26	3	14.3	12	1	A53709	alpha-conotoxin Im
27	3	14.3	12	2	S67528	napin - rape (frag
28	3	14.3	12	2	JU0356	cycloleoneurin
29	3	14.3	12	2	PN0162	malate dehydrogena
30	3	14.3	12	2	PH1188	T-cell receptor al
31	3	14.3	13	2	PH1772	T cell receptor al
32	3	14.3	13	2	B47415	mannose-1-phosphat
33	3	14.3	14	2	C40944	hypothetical prote
34	3	14.3	14	2	I39753	nitrogenase (EC 1.
35	3	14.3	14	2	A60737	pollen allergen Lo
36	3	14.3	14	2	PS0258	38K protein 3228 -
37	3	14.3	14	2	S57574	T cell receptor V-
38	3	14.3	14	2	S47366	T-cell antigen rec
39	3	14.3	14	2	S58426	spermadhesin AWN h
40	3	14.3	14	2	A43847	Ig mu chain V regi
41	3	14.3	14	2	PH0747	T-cell receptor be
42	3	14.3	15	2	I52734	gene c-Ki-ras prot
43	3	14.3	15	2	I49420	placental lactogen
44	3	14.3	15	2	S26528	T-cell receptor al
45	3	14.3	15	2	S36893	ribosomal protein

# ALIGNMENTS

## RESULT 1

B36042

oxfX protein - Escherichia coli plasmid RP4

C;Species: Escherichia coli

C;Date: 30-Nov-1990 #sequence\_revision 30-Nov-1990 #text\_change 23-Mar-1993

C;Accession: B36042

R;Pansegrau, W.; Balzer, D.; Kruff, V.; Lurz, R.; Lanka, E.

Proc. Natl. Acad. Sci. U.S.A. 87, 6555-6559, 1990

A;Title: In vitro assembly of relaxosomes at the transfer origin of plasmid RP4.

A;Reference number: A36042; MUID:90370824; PMID:2168553

A;Accession: B36042

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-13 <PAN>

C;Genetics:

A;Genome: plasmid

Query Match 23.8%; Score 5; DB 2; Length 13;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGGCA 14  
    ||||  
Db 9 AGGCA 13

RESULT 2

S16144

spoVA protein - Bacillus licheniformis (fragment)

C;Species: Bacillus licheniformis

C;Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 15-Oct-1999

C;Accession: S16144

R;Moldover, B.; Piggot, P.J.; Yudkin, M.D.

J. Gen. Microbiol. 137, 527-531, 1991

A;Title: Identification of the promoter and the transcriptional start site of the spoVA operon of Bacillus subtilis and Bacillus licheniformis.

A;Reference number: S16144; MUID:91237317; PMID:1903432

A;Accession: S16144

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-17 <MOL>

A;Cross-references: EMBL:X53991; NID:g39580; PIDN:CAA37938.1; PID:g39581

Query Match 19.0%; Score 4; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 RHRL 21  
    ||||  
Db 10 RHRL 13

RESULT 3

S46418

NTL1 protein - curled-leaved tobacco

C;Species: Nicotiana glauca (curled-leaved tobacco)

C;Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 02-Aug-2002

C;Accession: S46418

R;Daniel-Vedele, F.; Caboche, M.

Mol. Gen. Genet. 240, 365-373, 1993

A;Title: A tobacco cDNA clone encoding a GATA-1 zinc finger protein homologous to regulators of nitrogen metabolism in fungi.

A;Reference number: S46418; MUID:94019241; PMID:8413186

A;Accession: S46418

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-18 <DAN>

A;Cross-references: EMBL:S65745

C;Superfamily: Arabidopsis thaliana GATA transcription factor 4

Query Match 19.0%; Score 4; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 REGP 9  
    ||||  
Db 9 REGP 12

# RESULT 4

S19914

choline O-acetyltransferase (EC 2.3.1.6) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 05-May-2000

C;Accession: S19914

R;Toussaint, J.L.; Bausero, P.; Stricker, C.; Geoffroy, V.; Simoni, P.; Kempf, J.; Schmitt, M.

submitted to the EMBL Data Library, March 1992

A;Description: Human choline acetyltransferase gene: analysis of transcription control region and codon homology studies.

A;Reference number: S19914

A;Accession: S19914

A;Molecule type: DNA

A;Residues: 1-18 <TOU>

A;Cross-references: EMBL:X65023

C;Keywords: acyltransferase; coenzyme A

Query Match 19.0%; Score 4; DB 2; Length 18;  
Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GPAG 11  
| | | |  
Db 14 GPAG 17

# RESULT 5

B29501

fibrinopeptide A - European moose

C;Species: Alces alces alces (European moose, elk)

C;Date: 21-Nov-1987 #sequence\_revision 21-Nov-1987 #text\_change 18-Aug-2000

C;Accession: B29501

R;Blombaeck, B.; Blombaeck, M.; Hann, C.

unpublished results, cited by Blombaeck, B., and Blombaeck, M., in Chemotaxonomy and Serotaxonomy, Hawkes, J.G., ed., pp.3-20, Academic Press, London and New York, 1968

A;Reference number: A29501

A;Accession: B29501

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-19 <BL0>

C;Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disulfide ring homology

Query Match 19.0%; Score 4; DB 2; Length 19;  
Best Local Similarity 100.0%; Pred. No. 5.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 PAGG 12  
| | | |  
Db 6 PAGG 9

# RESULT 6

C39543

collagen alpha 3(IX) chain - bovine (fragment)  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 28-Feb-1992 #sequence\_revision 28-Feb-1992 #text\_change 23-May-1997  
C;Accession: C39543  
R;Wu, J.J.; Lark, M.W.; Chun, L.E.; Eyre, D.R.  
J. Biol. Chem. 266, 5625-5628, 1991  
A;Title: Sites of stromelysin cleavage in collagen types II, IX, X, and XI of cartilage.  
A;Reference number: A39543; MUID:91170231; PMID:2005102  
A;Accession: C39543  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-21 <WUA>  
C;Superfamily: unassigned collagens

Query Match 19.0%; Score 4; DB 2; Length 21;  
Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GPAG 11  
|||  
Db 14 GPAG 17

#### RESULT 7

PT0666

T-cell receptor beta chain V-D-J region (121-3BD) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C;Accession: PT0666

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Reference number: PT0509; MUID:91277601; PMID:1711558

A;Accession: PT0666

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-7 <FEE>

A;Experimental source: day 4 postnatal thymus, strain BALB/c

C;Keywords: T-cell receptor

Query Match 14.3%; Score 3; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGG 12  
|||  
Db 5 AGG 7

#### RESULT 8

PT0683

T-cell receptor beta chain V-D-J region (140-1N) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C;Accession: PT0683

R;Feeney, A.J.  
J. Exp. Med. 174, 115-124, 1991  
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A;Reference number: PT0509; MUID:91277601; PMID:1711558  
A;Accession: PT0683  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-7 <FEE>  
A;Experimental source: day 18 fetal thymus, strain BALB/c  
C;Keywords: T-cell receptor

Query Match 14.3%; Score 3; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGG 12  
|||  
Db 5 AGG 7

RESULT 9

PT0527

T-cell receptor beta chain V-D-J region (100-4I) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C;Accession: PT0527

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.

A;Reference number: PT0509; MUID:91277601; PMID:1711558

A;Accession: PT0527

A;Status: translation not shown

A;Molecule type: mRNA

A;Residues: 1-8 <FEE>

A;Experimental source: adult thymus, strain BALB/c

C;Keywords: T-cell receptor

Query Match 14.3%; Score 3; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGG 12  
|||  
Db 5 AGG 7

RESULT 10

PT0509

T-cell receptor beta chain V-D-J region (100-4N) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 30-May-1997

C;Accession: PT0509

R;Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.  
A;Reference number: PT0509; MUID:91277601; PMID:1711558  
A;Accession: PT0509  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-8 <FEE>  
A;Experimental source: adult thymus, strain BALB/c  
C;Keywords: T-cell receptor

Query Match 14.3%; Score 3; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGG 12  
|||  
Db 5 AGG 7

RESULT 11

I54017

granulocyte-colony stimulating factor precursor - synthetic (fragment)

C;Species: synthetic

A;Note: human gene engineered and expressed in Echerichia coli

C;Date: 28-Jan-2000 #sequence\_revision 28-Jan-2000 #text\_change 28-Jan-2000

C;Accession: I54017

R;Devlin, P.E.; Drummond, R.J.; Toy, P.; Mark, D.F.; Watt, K.W.; Devlin, J.J.  
Gene 65, 13-22, 1988

A;Title: Alteration of amino-terminal codons of human granulocyte-colony-stimulating factor increases expression levels and allows efficient processing by methionine aminopeptidase in Escherichia coli.

A;Reference number: I54017; MUID:88284374; PMID:2456256

A;Accession: I54017

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-8 <DEV>

A;Cross-references: GB:M20922; NID:g806638; PIDN:AAA66353.1; PID:g183043

Query Match 14.3%; Score 3; DB 4; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GPA 10  
|||  
Db 5 GPA 7

RESULT 12

S35538

ribosomal protein S7 - Mycobacterium intracellulare (fragment)

C;Species: Mycobacterium intracellulare

C;Date: 09-Dec-1993 #sequence\_revision 14-Nov-1997 #text\_change 13-Aug-1999

C;Accession: S35538

R;Nair, J.; Rouse, D.; Morris, S.

Nucleic Acids Res. 21, 1039, 1993

A;Title: Nucleotide sequence analysis of the ribosomal S12 gene of Mycobacterium intracellulare.

A;Reference number: S35537; MUID:93197130; PMID:8451173  
A;Accession: S35538  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-9 <NAI>  
A;Cross-references: EMBL:L08171; NID:g149994; PIDN:AAA25376.1; PID:g551901  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1993  
C;Superfamily: Escherichia coli ribosomal protein S7  
C;Keywords: protein biosynthesis; ribosome

Query Match 14.3%; Score 3; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GPA 10  
|||  
Db 5 GPA 7

#### RESULT 13

##### QDRB

delta sleep-inducing peptide - rabbit

C;Species: Oryctolagus cuniculus (domestic rabbit)

C;Date: 20-Jun-2000 #sequence\_revision 20-Jun-2000 #text\_change 20-Jun-2000

C;Accession: A01422

R;Monnier, M.; Dudler, L.; Gachter, R.; Maier, P.F.; Tobler, H.J.;

Schoenenberger, G.A.

Experientia 33, 548-552, 1977

A;Title: The delta sleep inducing peptide (DSIP). Comparative properties of the original and synthetic nonapeptide.

A;Reference number: A01422; MUID:77185324; PMID:862769

A;Accession: A01422

A;Molecule type: protein

A;Residues: 1-9 <MON>

C;Comment: This peptide was obtained from dialysates of occipital venous sinus blood from rabbits kept asleep by electric stimulation of the thalamus. When infused into the mesodiencephalic ventricle of recipient rabbits, it induces spindle and delta EEG activity and reduced motor activities.

C;Superfamily: unassigned animal peptides

Query Match 14.3%; Score 3; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGG 12  
|||  
Db 2 AGG 4

#### RESULT 14

##### A53797

3',5'-cyclic-GMP phosphodiesterase (EC 3.1.4.35) gamma chain - bullfrog (fragment)

C;Species: Rana catesbeiana (bullfrog)

C;Date: 01-Dec-1995 #sequence\_revision 01-Dec-1995 #text\_change 17-Mar-1999

C;Accession: A53797

R;Tsuboi, S.; Matsumoto, H.; Jackson, K.W.; Tsujimoto, K.; Williams, T.; Yamazaki, A.  
J. Biol. Chem. 269, 15016-15023, 1994  
A;Title: Phosphorylation of an inhibitory subunit of cGMP phosphodiesterase in Rana catesbiana rod photoreceptors. I. Characterization of the phosphorylation.  
A;Reference number: A53797; MUID:94253058; PMID:8195138  
A;Accession: A53797  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-9 <TSU>  
C;Keywords: cGMP binding; phosphoric diester hydrolase

Query Match 14.3%; Score 3; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GPA 10  
|||  
Db 4 GPA 6

RESULT 15

S78420  
ribosomal protein RL41, mitochondrial [validated] - rat (fragment)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 25-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 21-Jul-2000  
C;Accession: S78420  
R;Goldschmidt-Reisin, S.; Graack, H.R.  
submitted to the Protein Sequence Database, February 1998  
A;Reference number: S78411  
A;Accession: S78420  
A;Molecule type: protein  
A;Residues: 1-9 <GOL>  
A;Note: the protein is designated as mitochondrial ribosomal protein L41  
C;Keywords: mitochondrion; protein biosynthesis; ribosome

Query Match 14.3%; Score 3; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 HRL 21  
|||  
Db 5 HRL 7

Search completed: November 13, 2003, 10:39:57  
Job time : 21.5625 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:55:06 ; Search time 11.375 Seconds  
(without alignments)  
86.819 Million cell updates/sec



Title: US-09-228-866-16  
 Perfect score: 21  
 Sequence: 1 WRCVLREGPAGGCAWFNRHRL 21  
 Scoring table: OLIGO  
 Gapop 60.0 , Gapext 60.0  
 Searched: 127863 seqs, 47026705 residues  
 Word size : 0

Total number of hits satisfying chosen parameters: 1246

Minimum DB seq length: 7  
 Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query		DB	ID	Description
	Score	Match Length			
1	4	19.0	16	1 FIBA_ODOHE	P14459 odocoileus
2	4	19.0	17	1 SP51_BACLI	P27642 bacillus li
3	4	19.0	19	1 FIBA_RANTA	P14462 rangifer ta
4	3	14.3	8	1 COW2_CONPU	P58785 conus purpu
5	3	14.3	8	1 RS7_MYCIT	P33564 mycobacteri
6	3	14.3	8	1 WP1_PERAT	P83195 perkinsus a
7	3	14.3	9	1 DSIP_RABIT	P01158 oryctolagus
8	3	14.3	10	1 COXO_RAT	P80432 rattus norv
9	3	14.3	10	1 COXO_THUOB	P80982 thunnus obe
10	3	14.3	11	1 NUHM_CANFA	P49820 canis famil
11	3	14.3	12	1 CXA1_CONIM	P50983 conus imper
12	3	14.3	13	1 FIBA_CAVPO	P14445 cavia porce
13	3	14.3	13	1 UN10_CLOPA	P81345 clostridium
14	3	14.3	15	1 CHI1_PEA	P21225 pisum sativ
15	3	14.3	15	1 TA1_TREBR	P34070 tremella br
16	3	14.3	16	1 AF1S_MALPA	P83140 malva parvi
17	3	14.3	17	1 PA2_AUSSU	P59066 austrelaps
18	3	14.3	18	1 GOME_ACAGO	P82358 acanthoscur
19	3	14.3	18	1 MLB_HORSE	P01202 equus cabal
20	3	14.3	19	1 CXR_CONTU	P58811 conus tulip
21	3	14.3	20	1 CD4_SHEEP	P05542 ovis aries
22	3	14.3	20	1 CISY_STRHY	P20903 streptomyce
23	3	14.3	20	1 OMPW_VIBAL	P83151 vibrio algi
24	3	14.3	20	1 PUFK_RHOSH	Q53121 rhodobacter
25	3	14.3	20	1 STYB_STYCL	P81470 styela clav
26	3	14.3	21	1 CSPS_STRTR	P81622 streptococc
27	3	14.3	21	1 CXKT_CONTU	P17684 conus tulip

28	3	14.3	21	1	HCY4_MAI SQ	P82305	maia squina
29	3	14.3	21	1	MDH_KLEPN	P80535	klebsiella
30	3	14.3	21	1	REV_HV2D2	P15830	human immun
31	2	9.5	7	1	BRHP_CONIM	P58803	conus imper
32	2	9.5	7	1	CARP_MYTED	P10420	mytilus edu
33	2	9.5	7	1	FAR1_HELTI	P41871	helisoma tr
34	2	9.5	7	1	FAR1_MACRS	P83274	macrobrachi
35	2	9.5	7	1	FAR1_PROCL	P38499	procambarus
36	2	9.5	7	1	FAR2_ASCSU	P31890	ascaris suu
37	2	9.5	7	1	FAR2_PROCL	P38498	procambarus
38	2	9.5	7	1	FAR5_HIRME	P42564	hirudo medi
39	2	9.5	7	1	IGAO_DACDE	P06294	dactylum d
40	2	9.5	7	1	MNP1_LEPDE	P42984	leptinotars
41	2	9.5	7	1	TPFY_PACDA	P83455	pachymedusa
42	2	9.5	7	1	TY51_LITRU	P82065	litoria rub
43	2	9.5	7	1	UH11_RAT	P56576	rattus norv
44	2	9.5	7	1	WWA1_ACHFU	P35919	achatina fu
45	2	9.5	8	1	AL15_CARMA	P81818	carcinus ma

# ALIGNMENTS

## RESULT 1

### FIBA\_ODOHE

ID FIBA\_ODOHE STANDARD; PRT; 16 AA.

AC P14459;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).

GN FGA.

OS Odocoileus hemionus (Mule deer) (Black-tailed deer).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;

OC Cervidae; Odocoileinae; Odocoileus.

OX NCBI\_TaxID=9872;

RN [1]

RP SEQUENCE.

RX MEDLINE=67209145; PubMed=6033721;

RA Doolittle R.F., Schubert D., Schwartz S.A.;

RT "Amino acid sequence studies on artiodactyl fibrinopeptides. I.

RT Dromedary camel, mule deer, and cape buffalo.";

RL Arch. Biochem. Biophys. 118:456-467(1967).

CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT

CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET

CC AGGREGATION.

CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS

CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.

CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY

CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA

CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES

CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.

KW Blood coagulation; Plasma.

FT PEPTIDE 1 16 FIBRINOPEPTIDE A.

FT NON\_TER 16 16

SQ SEQUENCE 16 AA; 1519 MW; E8098EB4B33B59C5 CRC64;

Query Match 19.0%; Score 4; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PAGG 12  
    ||||  
Db 3 PAGG 6

RESULT 2

SP51\_BACLI

ID SP51\_BACLI STANDARD; PRT; 17 AA.  
AC P27642;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE Stage V sporulation protein AA (Fragment).  
GN SPOVAA.  
OS Bacillus licheniformis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1402;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91237317; PubMed=1903432;  
RA Moldover B., Piggot P.J., Yudkin M.D.;  
RT "Identification of the promoter and the transcriptional start site of  
RT the spoVA operon of Bacillus subtilis and Bacillus licheniformis.";  
RL J. Gen. Microbiol. 137:527-531(1991).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X53991; CAA37938.1; -.  
DR PIR; S16144; S16144.  
KW Sporulation.  
FT NON\_TER 17 17  
SQ SEQUENCE 17 AA; 2167 MW; 17505DA08BFE133B CRC64;

Query Match 19.0%; Score 4; DB 1; Length 17;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 RHRL 21  
    ||||  
Db 10 RHRL 13

RESULT 3

FIBA\_RANTA

ID FIBA\_RANTA STANDARD; PRT; 19 AA.  
AC P14462;

DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).  
 GN FGA.  
 OS Rangifer tarandus (Reindeer) (Caribou).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervioidea;  
 OC Cervidae; Odocoileinae; Rangifer.  
 OX NCBI\_TaxID=9870;  
 RN [1]  
 RP SEQUENCE.  
 RA Blomback B., Blomback M., Grondahl N.J.;  
 RT "Studies on fibrinopeptides from mammals.";  
 RL Acta Chem. Scand. 19:1789-1791(1965).  
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT  
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET  
 CC AGGREGATION.  
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS  
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY  
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.  
 KW Blood coagulation; Plasma.  
 FT PEPTIDE 1 19 FIBRINOPEPTIDE A.  
 FT NON\_TER 19 19  
 SQ SEQUENCE 19 AA; 1762 MW; 9BA41873C33D9F05 CRC64;

Query Match 19.0%; Score 4; DB 1; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 PAGG 12  
 ||||  
 Db 6 PAGG 9

#### RESULT 4

##### COW2\_CONPU

ID COW2\_CONPU STANDARD; PRT; 8 AA.  
 AC P58785;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Leu-contryphan-P.  
 OS Conus purpurascens (Purple cone).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
 OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;  
 OC Neogastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=41690;  
 RN [1]  
 RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.  
 RC STRAIN=Clipperton Island; TISSUE=Venom;  
 RX MEDLINE=99388839; PubMed=10461743;  
 RA Jacobsen R.B., Jimenez E.C., De la Cruz R.G.C., Gray W.R., Cruz L.J.,  
 RA Olivera B.M.;

RT "A novel D-leucine-containing Conus peptide: diverse conformational  
 RT dynamics in the contryphan family.";  
 RL J. Pept. Res. 54:93-99(1999).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.  
 CC -!- MASS SPECTROMETRY: MW=888.4; METHOD=LSIMS.  
 CC -!- SIMILARITY: BELONGS TO THE CONTRYPHAN FAMILY.  
 KW Toxin; Hydroxylation; D-amino acid.  
 FT DISULFID 2 8  
 FT MOD\_RES 4 4 D-LEUCINE.  
 SQ SEQUENCE 8 AA; 890 MW; 75A367672732CEB8 CRC64;

Query Match 14.3%; Score 3; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CVL 5  
 |||  
 Db 2 CVL 4

# RESULT 5

## RS7\_MYCIT

ID RS7\_MYCIT STANDARD; PRT; 8 AA.  
 AC P33564;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 30S ribosomal protein S7 (Fragment).  
 GN RPSG.  
 OS Mycobacterium intracellulare.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1767;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93197130; PubMed=8451173;  
 RA Nair J., Rouse D.A., Morris S.L.;  
 RT "Nucleotide sequence analysis of the ribosomal S12 gene of  
 RT Mycobacterium intracellulare.";  
 RL Nucleic Acids Res. 21:1039-1039(1993).  
 CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds  
 CC directly to 16S rRNA where it nucleates assembly of the head  
 CC domain of the 30S subunit. Is located at the subunit interface  
 CC close to the decoding center, probably blocks exit of the E-site  
 CC tRNA (By similarity).  
 CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts proteins S9  
 CC and S11 (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE S7P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC -----
DR EMBL; L08171; AAA25376.1; -.
DR PIR; S35538; S35538.
DR HAMAP; MF_00480; -; 1.
DR InterPro; IPR000235; Ribosomal_S7.
DR PROSITE; PS00052; RIBOSOMAL_S7; PARTIAL.
KW Ribosomal protein; RNA-binding; rRNA-binding; tRNA-binding.
FT INIT_MET      0      0      BY SIMILARITY.
FT NON_TER       8      8
SQ SEQUENCE      8 AA;  850 MW;  63276DC768732417 CRC64;

```

```

Query Match          14.3%;  Score 3;  DB 1;  Length 8;
Best Local Similarity 100.0%;  Pred. No. 1.3e+05;
Matches      3;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

```

```

Qy      8 GPA 10
      |||
Db      4 GPA 6

```

# RESULT 6

## WP1\_PERAT

```

ID WP1_PERAT      STANDARD;      PRT;      8 AA.
AC P83195;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Wall protein-1 (PWP-1) (Fragment).
OS Perkinsus atlanticus.
OC Eukaryota; Alveolata; Perkinsea; Perkinsida; Perkinsidae; Perkinsus.
OX NCBI_TaxID=106964;
RN [1]
RP SEQUENCE, FUNCTION, SUBCELLULAR LOCATION, AND DEVELOPMENTAL STAGE.
RX MEDLINE=22044350; PubMed=12049410;
RA Montes J.F., Durfort M., Llado A., Garcia-Valero J.;
RT "Characterization and immunolocalization of a main proteinaceous
RT component of the cell wall of the protozoan parasite Perkinsus
RT atlanticus.";
RL Parasitology 124:477-484(2002).
CC -!- FUNCTION: Is a major protein component of the cell wall. May play
CC      a key role in the organization of the cell wall and in promoting
CC      the survival of this parasite.
CC -!- SUBCELLULAR LOCATION: Cell wall. Disulfide-linked to other cell
CC      wall components.
CC -!- DEVELOPMENTAL STAGE: Expressed throughout all walled developmental
CC      stages.
KW Cell wall.
FT NON_TER       8      8
SQ SEQUENCE      8 AA;  765 MW;  F1787DD87B1AAB16 CRC64;

```

```

Query Match          14.3%;  Score 3;  DB 1;  Length 8;
Best Local Similarity 100.0%;  Pred. No. 1.3e+05;
Matches      3;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

```

```

Qy      10 AGG 12
      |||
Db      6 AGG 8

```

RESULT 7

DSIP\_RABIT

ID DSIP\_RABIT STANDARD; PRT; 9 AA.  
AC P01158;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Delta sleep-inducing peptide (DSIP).  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=77185324; PubMed=862769;  
RA Monnier M., Dudler L., Gachter R., Maier P.F., Tobler H.J.,  
RA Schoenenberger G.A.;  
RT "The delta sleep inducing peptide (DSIP). Comparative properties of  
RT the original and synthetic nonapeptide."  
RL Experientia 33:548-552(1977).  
RN [2]  
RP SEQUENCE, AND SYNTHESIS.  
RX MEDLINE=79054421; PubMed=568769;  
RA Schoenenberger G.A., Maier P.F., Tobler H.J., Wilson K., Monnier M.;  
RT "The delta EEG (sleep)-inducing peptide (DSIP). XI. Amino-acid  
RT analysis, sequence, synthesis and activity of the nonapeptide."  
RL Pflugers Arch. 376:119-129(1978).  
RN [3]  
RP REVIEW.  
RX MEDLINE=87175129; PubMed=3550726;  
RA Graf M.V., Kastin A.J.;  
RT "Delta-sleep-inducing peptide (DSIP): an update."  
RL Peptides 7:1165-1187(1986).  
CC -!- FUNCTION: WHEN INFUSED INTO THE MESODIENCEPHALIC VENTRICLE OF  
CC RECIPIENT RABBITS INDUCES SPINDLE AND DELTA EEG ACTIVITY AND  
CC REDUCED MOTOR ACTIVITIES.  
CC -!- MISCELLANEOUS: THIS PEPTIDE WAS OBTAINED FROM DIALYSATES OF  
CC OCCIPITAL VENOUS SINUS BLOOD FROM RABBITS KEPT ASLEEP BY ELECTRIC  
CC STIMULATION OF THE THALAMUS.  
CC -!- DATABASE: NAME=Protein Spotlight;  
CC NOTE=Issue 8 of March 2001;  
CC WWW="http://www.expasy.org/spotlight/articles/sptlt008.html".  
DR PIR; A01422; QDRB.  
SQ SEQUENCE 9 AA; 849 MW; DDD365BDDAA8787D CRC64;

Query Match 14.3%; Score 3; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 AGG 12  
|||  
Db 2 AGG 4

RESULT 8

COXO\_RAT

ID COXO\_RAT STANDARD; PRT; 10 AA.

AC P80432;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Cytochrome c oxidase polypeptide VIIC, mitochondrial (EC 1.9.3.1)

DE (VIIIA) (Fragment).

GN COX7C OR COX7C1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE.

RC STRAIN=Wistar; TISSUE=Heart, and Liver;

RX MEDLINE=95324529; PubMed=7601105;

RA Schaeffer H., Noack H., Halangk W., Brandt U., von Jagow G.;

RT "Cytochrome-c oxidase in developing rat heart. Enzymic properties and

RT amino-terminal sequences suggest identity of the fetal heart and the

RT adult liver isoform.";

RL Eur. J. Biochem. 230:235-241(1995).

CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE

CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN

CC MITOCHONDRIAL ELECTRON TRANSPORT.

CC -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome

CC c + 2 H(2)O.

CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIC FAMILY.

DR PIR; S65388; S65388.

KW Oxidoreductase; Mitochondrion.

FT NON\_TER 10 10

SQ SEQUENCE 10 AA; 1117 MW; 126DE767687B1B0B CRC64;

Query Match 14.3%; Score 3; DB 1; Length 10;

Best Local Similarity 100.0%; Pred. No. 1.6e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 EGP 9

Db 5 EGP 7

# RESULT 9

COXO\_THUOB

ID COXO\_THUOB STANDARD; PRT; 10 AA.

AC P80982;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Cytochrome c oxidase polypeptide VIIC (EC 1.9.3.1) (Fragment).

OS Thunnus obesus (Bigeye tuna).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;

OC Scombridae; Thunnus.

OX NCBI\_TaxID=8241;

RN [1]



RP SEQUENCE.  
 RC TISSUE=Heart, and Liver;  
 RX MEDLINE=97454291; PubMed=9310366;  
 RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,  
 RA Kadenbach B.;  
 RT "The subunit structure of cytochrome-c oxidase from tuna heart and  
 RT liver.";  
 RL Eur. J. Biochem. 248:99-103(1997).  
 CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE  
 CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN  
 CC MITOCHONDRIAL ELECTRON TRANSPORT.  
 CC -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome  
 CC c + 2 H(2)O.  
 CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIC FAMILY.  
 DR PIR; S77990; S77990.  
 KW Oxidoreductase; Inner membrane; Mitochondrion.  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1059 MW; 126DE767687B1DCB CRC64;

Query Match 14.3%; Score 3; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 EGP 9  
 |||  
 Db 5 EGP 7

# RESULT 10

## NUHM\_CANFA

ID NUHM\_CANFA STANDARD; PRT; 11 AA.  
 AC P49820;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE NADH-ubiquinone oxidoreductase 24 kDa subunit (EC 1.6.5.3)  
 DE (EC 1.6.99.3) (Fragment).  
 GN NDUFV2.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Heart;  
 RX MEDLINE=98163340; PubMed=9504812;  
 RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
 RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
 RT dog heart proteins.";  
 RL Electrophoresis 18:2795-2802(1997).  
 CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY  
 CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED  
 CC TO BE UBIQUINONE. COMPONENT OF THE FLAVOPROTEIN-SULFUR (FP)  
 CC FRAGMENT OF THE ENZYME.  
 CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
 CC -!- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.

CC -!- COFACTOR: Binds 1 2Fe-2S cluster (Potential).  
 CC -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.  
 CC THIS IS A COMPONENT OF THE FLAVOPROTEIN FRACTION.  
 CC -!- SUBCELLULAR LOCATION: Matrix and cytoplasmic side of the  
 CC mitochondrial inner membrane.  
 CC -!- SIMILARITY: BELONGS TO THE COMPLEX I 24 kDa SUBUNIT FAMILY.  
 DR HSC-2DPAGE; P49820; DOG.  
 DR InterPro; IPR002023; Cmplx1\_24kDa.  
 DR PROSITE; PS01099; COMPLEX1\_24K; PARTIAL.  
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Metal-binding;  
 KW Iron-sulfur; Iron; 2Fe-2S.  
 FT NON\_TER 11 11  
 SQ SEQUENCE 11 AA; 1099 MW; 267F5369C9C72DD8 CRC64;

Query Match 14.3%; Score 3; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGG 12  
 |||  
 Db 2 AGG 4

# RESULT 11

## CXA1\_CONIM

ID CXA1\_CONIM STANDARD; PRT; 12 AA.  
 AC P50983;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Alpha-conotoxin ImI.  
 OS Conus imperialis (Imperial cone).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;  
 OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;  
 OC Neogastropoda; Conoidea; Conidae; Conus.  
 OX NCBI\_TaxID=35631;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Venom;  
 RX MEDLINE=94266889; PubMed=8206995;  
 RA McIntosh J.M., Yoshikami D., Mahe E., Nielsen D.B., Rivier J.E.,  
 RA Gray W.R., Olivera B.M.;  
 RT "A nicotinic acetylcholine receptor ligand of unique specificity,  
 RT alpha-conotoxin ImI.";  
 RL J. Biol. Chem. 269:16733-16739(1994).  
 RN [2]  
 RP CHARACTERIZATION.  
 RX MEDLINE=95379776; PubMed=7651351;  
 RA Johnson D.S., Martinez J., Elgoyhen A.B., Heinemann S.F.,  
 RA McIntosh J.M.;  
 RT "Alpha-conotoxin ImI exhibits subtype-specific nicotinic  
 RT acetylcholine receptor blockade: preferential inhibition of homomeric  
 RT alpha 7 and alpha 9 receptors.";  
 RL Mol. Pharmacol. 48:194-199(1995).  
 RN [3]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=99212205; PubMed=10194298;

RA Rogers J.P., Luginbuehl P., Shen G.S., McCabe R.T., Stevens R.C.,  
 RA Wemmer D.E.;  
 RT "NMR solution structure of alpha-conotoxin ImI and comparison to  
 RT other conotoxins specific for neuronal nicotinic acetylcholine  
 RT receptors.";  
 RL Biochemistry 38:3874-3882(1999).  
 RN [4]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=99280313; PubMed=10350614;  
 RA Gouda H., Hirono S.;  
 RT "Solution structure of alpha-conotoxin ImI determined by  
 RT two-dimensional NMR spectroscopy.";  
 RL Biochim. Biophys. Acta 1431:384-394(1999).  
 RN [5]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=99158061; PubMed=10050774;  
 RA Maslennikov I.V., Shenkarev Z.O., Zhmak M.N., Ivanov V.T.,  
 RA Methfessel C., Tsetlin V.I., Arseniev A.S.;  
 RT "NMR spatial structure of alpha-conotoxin ImI reveals a common  
 RT scaffold in snail and snake toxins recognizing neuronal nicotinic  
 RT acetylcholine receptors.";  
 RL FEBS Lett. 444:275-280(1999).  
 RN [6]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=99358772; PubMed=10431825;  
 RA Lamthanh H., Jegou-Matheron C., Servent D., Menez A., Lancelin J.-M.;  
 RT "Minimal conformation of the alpha-conotoxin ImI for the alpha7  
 RT neuronal nicotinic acetylcholine receptor recognition: correlated CD,  
 RT NMR and binding studies.";  
 RL FEBS Lett. 454:293-298(1999).  
 RN [7]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=99324017; PubMed=10395477;  
 RA Gehrman J., Daly N.L., Alewood P.F., Craik D.J.;  
 RT "Solution structure of alpha-conotoxin ImI by 1H nuclear magnetic  
 RT resonance.";  
 RL J. Med. Chem. 42:2364-2372(1999).  
 RN [8]  
 RP MUTAGENESIS OF ASP-5; ARG-7 AND ARG-11, AND STRUCTURE BY NMR OF THESE  
 RP THREE MUTANTS.  
 RX MEDLINE=20574023; PubMed=11124036;  
 RA Rogers J.P., Luginbuehl P., Pemberton K., Harty P., Wemmer D.E.,  
 RA Stevens R.C.;  
 RT "Structure-activity relationships in a peptidic alpha7 nicotinic  
 RT acetylcholine receptor antagonist.";  
 RL J. Mol. Biol. 304:911-926(2000).  
 CC -!- FUNCTION: Alpha-conotoxins act on postsynaptic membranes, they  
 CC bind to the nicotinic acetylcholine receptors (nAChR) and thus  
 CC inhibit them. It is highly active against the neuromuscular  
 CC receptor in frog but not in mice. In contrast, it induces seizures  
 CC when injected centrally in mice and rats. It targets neuronal  
 CC nAChRs in mammals. It blocks homomeric alpha-7 nicotinic receptors  
 CC with the highest apparent affinity and homomeric alpha-9 receptors  
 CC with 8-fold lower affinity. It has no effect on receptors composed  
 CC of alpha-2/beta-2, alpha-3/beta-2, alpha-4/beta-2, alpha-2/beta-4,  
 CC alpha-3/beta-4, or alpha-4/beta-4 subunit combinations.  
 CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.  
 CC -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE  
 CC FAMILY.  
 DR PIR; A53709; A53709.  
 DR PDB; 1IM1; 15-JUN-99.  
 DR PDB; 1IMI; 23-APR-99.  
 DR PDB; 1CNL; 27-MAY-99.  
 DR PDB; 1E74; 27-DEC-00.  
 DR PDB; 1E75; 27-DEC-00.  
 DR PDB; 1E76; 27-DEC-00.  
 DR PDB; 1G2G; 08-NOV-00.  
 KW Postsynaptic neurotoxin; Neurotoxin; Toxin;  
 KW Acetylcholine receptor inhibitor; Amidation; 3D-structure.  
 FT DISULFID 2 8  
 FT DISULFID 3 12  
 FT MOD\_RES 12 12 AMIDATION.  
 FT MUTAGEN 5 5 D->L: REDUCTION OF TOXICITY.  
 FT MUTAGEN 7 7 R->L: REDUCTION OF TOXICITY.  
 FT MUTAGEN 11 11 R->E: NO LOSS OF ACTIVITY.  
 FT HELIX 2 4  
 FT HELIX 6 8  
 FT TURN 10 12  
 SQ SEQUENCE 12 AA; 1357 MW; 9C29CEA545A4176A CRC64;

Query Match 14.3%; Score 3; DB 1; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WRC 3  
 |||  
 Db 10 WRC 12

# RESULT 12

FIBA\_CAVPO

ID FIBA\_CAVPO STANDARD; PRT; 13 AA.

AC P14445;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-JAN-1990 (Rel. 13, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).

GN FGA.

OS Cavia porcellus (Guinea pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.

OX NCBI\_TaxID=10141;

RN [1]

RP SEQUENCE.

RA Blomback B., Blomback M., Grondahl N.J.;

RT "Studies on fibrinopeptides from mammals.";

RL Acta Chem. Scand. 19:1789-1791(1965).

CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT

CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET

CC AGGREGATION.

CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS

CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.

CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY

CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA  
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES  
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.

KW Blood coagulation; Plasma.

FT PEPTIDE 1 13 FIBRINOPEPTIDE A.

FT NON\_TER 13 13

SQ SEQUENCE 13 AA; 1309 MW; 639999286C79DDDB CRC64;

Query Match 14.3%; Score 3; DB 1; Length 13;

Best Local Similarity 100.0%; Pred. No. 2e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGG 12

|||

Db 8 AGG 10

#### RESULT 13

UN10\_CLOPA

ID UN10\_CLOPA STANDARD; PRT; 13 AA.

AC P81345;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE Unknown protein CP 10 from 2D-page (Fragment).

OS Clostridium pasteurianum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI\_TaxID=1501;

RN [1]

RP SEQUENCE.

RC STRAIN=W5;

RX MEDLINE=98291870; PubMed=9629918;

RA Flengsrud R., Skjeldal L.;

RT "Two-dimensional gel electrophoresis separation and N-terminal

RT sequence analysis of proteins from Clostridium pasteurianum W5.";

RL Electrophoresis 19:802-806(1998).

CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN

CC PROTEIN IS: 5.2, ITS MW IS: 52.5 kDa.

FT NON\_TER 13 13

SQ SEQUENCE 13 AA; 1447 MW; AEBE730C4B90C056 CRC64;

Query Match 14.3%; Score 3; DB 1; Length 13;

Best Local Similarity 100.0%; Pred. No. 2e+03;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GPA 10

|||

Db 11 GPA 13

#### RESULT 14

CHI1\_PEA

ID CHI1\_PEA STANDARD; PRT; 15 AA.

AC P21225;

DT 01-MAY-1991 (Rel. 18, Created)

DT 01-MAY-1991 (Rel. 18, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Endochitinase A1 (EC 3.2.1.14) (Fragment).  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.  
 OX NCBI\_TaxID=3888;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=cv. Birte; TISSUE=Leaf;  
 RA Vad K., Mikkelsen J.D., Collinge D.B.;  
 RT "Induction, purification and characterization of chitinase isolated  
 RT from pea leaves inoculated with Ascochyta pisi.";  
 RL Planta 184:24-29(1991).  
 CC -!- FUNCTION: THIS PROTEIN FUNCTIONS AS A DEFENSE AGAINST CHITIN  
 CC CONTAINING FUNGAL PATHOGENS.  
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-  
 CC acetyl-D-glucosamine polymers of chitin.  
 CC -!- INDUCTION: By infection with the fungal pathogen Ascochyta pisi.  
 CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS IA WHICH IS SIMILAR TO  
 CC CLASS IB IN THE CATALYTIC PART BUT WHICH INCLUDE A N-TERMINAL  
 CC CHITIN-BINDING LECTIN DOMAIN (BELONGS TO FAMILY 19 OF GLYCOSYL  
 CC HYDROLASES).  
 DR InterPro; IPR001002; Chitin\_binding\_1.  
 DR InterPro; IPR000726; Glyco\_hydro\_19.  
 DR PROSITE; PS00026; CHITIN\_BINDING; PARTIAL.  
 DR PROSITE; PS00773; CHITINASE\_19\_1; PARTIAL.  
 DR PROSITE; PS00774; CHITINASE\_19\_2; PARTIAL.  
 KW Hydrolase; Glycosidase; Chitin degradation; Chitin-binding;  
 KW Multigene family.  
 FT NON\_TER 15 15  
 SQ SEQUENCE 15 AA; 1438 MW; 23355ED6BB11E869 CRC64;

Query Match 14.3%; Score 3; DB 1; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+03;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 AGG 12  
 |||  
 Db 7 AGG 9

# RESULT 15

## TA1\_TREBR

ID TA1\_TREBR STANDARD; PRT; 15 AA.  
 AC P34070;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE Tremmerogen A-I.  
 OS Tremella brasiliensis (Jelly fungus).  
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;  
 OC Tremellomycetidae; Tremellales; Tremellaceae; Tremella.  
 OX NCBI\_TaxID=29896;  
 RN [1]  
 RP SEQUENCE.  
 RA Ishibashi Y., Sakagami Y., Isogai A., Suzuki A.;

RT "Structures of Tremmerogens A-9291-I and A-9291-VIII: peptidyl sex  
RT hormones of Tremella brasiliensis.";  
RL Biochemistry 23:1399-1404(1984).  
CC -!- FUNCTION: TREMEROGEN A-I IS PRODUCED BY THE A MATING-TYPE CELLS  
CC AND INDUCES FORMATION OF CONJUGATION TUBES IN A MATING-TYPE CELLS.  
KW Pheromone; Prenylation; Lipoprotein.  
FT LIPID 15 15 FARNESYL.  
SQ SEQUENCE 15 AA; 1339 MW; 3AABA4FC2D605333 CRC64;

Query Match 14.3%; Score 3; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.3e+03;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 GGC 13  
|||  
Db 13 GGC 15

Search completed: November 13, 2003, 10:34:01  
Job time : 12.375 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:58:36 ; Search time 51.625 Seconds  
(without alignments)  
104.971 Million cell updates/sec

Title: US-09-228-866-16  
Perfect score: 21  
Sequence: 1 WRCVLREGPAGGCAWFNRHRL 21

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 7516

Minimum DB seq length: 7  
Maximum DB seq length: 21

Post-processing: Listing first 45 summaries

Database : SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*

```

8:  sp_organelle:*
9:  sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	5	23.8	13	2	Q47607	Q47607 escherichia
2	5	23.8	16	5	Q9TWK0	Q9twk0 mytilus edu
3	5	23.8	21	12	Q9JG25	Q9jg25 tt virus. o
4	4	19.0	11	12	Q83083	Q83083 leucania se
5	4	19.0	12	8	O03816	O03816 metasequoia
6	4	19.0	12	8	O03815	O03815 abies alba
7	4	19.0	12	8	P92454	P92454 cycas revol
8	4	19.0	12	8	Q36669	Q36669 pinus sylve
9	4	19.0	12	8	Q37790	Q37790 larix eurol
10	4	19.0	13	8	P92460	P92460 taxus bacca
11	4	19.0	13	8	Q36622	Q36622 picea abies
12	4	19.0	15	4	Q9UC90	Q9uc90 homo sapien
13	4	19.0	16	4	Q9UC88	Q9uc88 homo sapien
14	4	19.0	16	4	Q9UC54	Q9uc54 homo sapien
15	4	19.0	16	5	Q9TWK1	Q9twk1 mytilus edu
16	4	19.0	18	4	Q9UE43	Q9ue43 homo sapien
17	4	19.0	19	4	Q9UCH5	Q9uch5 homo sapien
18	4	19.0	19	6	Q8MJ41	Q8mj41 bos taurus
19	4	19.0	20	11	Q61871	Q61871 mus musculu
20	4	19.0	21	2	Q9R4X6	Q9r4x6 eubacterium
21	3	14.3	8	4	Q9Y4X6	Q9y4x6 homo sapien
22	3	14.3	8	12	O89965	O89965 polyomaviru
23	3	14.3	8	12	Q83349	Q83349 murine coro
24	3	14.3	8	12	Q84156	Q84156 orf virus (
25	3	14.3	9	4	Q9H326	Q9h326 homo sapien
26	3	14.3	9	11	Q99JF4	Q99jf4 mus musculu
27	3	14.3	9	12	P90359	P90359 barley mild
28	3	14.3	9	12	Q9IBM8	Q9ibm8 simian viru
29	3	14.3	9	12	Q9PYK1	Q9pyk1 simian viru
30	3	14.3	10	4	O60912	O60912 homo sapien
31	3	14.3	10	4	Q14096	Q14096 homo sapien
32	3	14.3	10	4	Q9H1I5	Q9h1i5 homo sapien
33	3	14.3	10	4	Q8N6B1	Q8n6b1 homo sapien
34	3	14.3	10	5	P82383	P82383 drosophila
35	3	14.3	10	8	Q8W7U4	Q8w7u4 anolis nite
36	3	14.3	10	10	P82938	P82938 hordeum vul
37	3	14.3	10	11	Q9QVF0	Q9qvf0 mus sp. pro



38	3	14.3	10	11	Q63389	Q63389 rattus norv
39	3	14.3	10	12	P90373	P90373 pseudorabie
40	3	14.3	11	2	Q47567	Q47567 escherichia
41	3	14.3	11	3	Q9HFN8	Q9hfn8 candida rug
42	3	14.3	11	4	Q9UNL8	Q9unl8 homo sapien
43	3	14.3	11	4	Q9C057	Q9c057 homo sapien
44	3	14.3	11	4	Q9UC46	Q9uc46 homo sapien
45	3	14.3	11	13	Q9PS64	Q9ps64 gallus gall

# ALIGNMENTS

## RESULT 1

Q47607

ID Q47607 PRELIMINARY; PRT; 13 AA.  
AC Q47607;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Coli plasmid RP4 traF (5'end), traG, traH, traI, traJ, traK, traL and  
DE traM genes of the transfer region.  
OS Escherichia coli.  
OG Plasmid RP4.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HB101;  
RX MEDLINE=92190548; PubMed=1665997;  
RA Ziegelin G., Pansegrau W., Strack B., Balzer D., Kroger M., Kruff V.,  
RA Lanka E.;  
RT "Nucleotide sequence and organization of genes flanking the transfer  
RT origin of promiscuous plasmid RP4.";  
RL DNA Seq. 1:303-327(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HB101;  
RX MEDLINE=92381045; PubMed=1324929;  
RA Ziegelin G., Pansegrau W., Lurz R., Lanka E.;  
RT "TraK protein of conjugative plasmid RP4 forms a specialized  
RT nucleoprotein complex with the transfer origin.";  
RL J. Biol. Chem. 267:17279-17286(1992).  
DR EMBL; X54459; CAA38337.1; -.  
KW Plasmid.  
SQ SEQUENCE 13 AA; 1241 MW; C88592E626F8A33D CRC64;

Query Match 23.8%; Score 5; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	10	AGGCA	14
Db	9	AGGCA	13

RESULT 2

Q9TWK0

ID Q9TWK0 PRELIMINARY; PRT; 16 AA.  
 AC Q9TWK0;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE Proximal collagen (Fragment).  
 OS Mytilus edulis (Blue mussel).  
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Mytiloida;  
 OC Mytiloidea; Mytilidae; Mytilus.  
 OX NCBI\_TaxID=6550;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95230211; PubMed=7714453;  
 RA Qin X., Waite J.H.;  
 RT "Exotic collagen gradients in the byssus of the mussel Mytilus  
 edulis."  
 RL J. Exp. Biol. 198:633-644(1995).  
 SQ SEQUENCE 16 AA; 1347 MW; B57F18783A18783C CRC64;

Query Match 23.8%; Score 5; DB 5; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 EGPAG 11  
 |||||  
 Db 9 EGPAG 13

RESULT 3

Q9JG25

ID Q9JG25 PRELIMINARY; PRT; 21 AA.  
 AC Q9JG25;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE ORF2 (Fragment).  
 OS TT virus.  
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.  
 OX NCBI\_TaxID=68887;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20261737; PubMed=10799591;  
 RA Okamoto H., Ukita M., Nishizawa T., Kishimoto J., Hoshi Y., Mizuo H.,  
 RA Tanaka T., Miyakawa Y., Mayumi M.;  
 RT "Circular double-stranded forms of TT virus DNA in the liver."  
 RL J. Virol. 74:5161-5167(2000).  
 DR EMBL; AB040782; BAA94203.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 21 21  
 SQ SEQUENCE 21 AA; 1949 MW; 39AD5337BFACABEC CRC64;

Query Match 23.8%; Score 5; DB 12; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GPAGG 12  
    ||||  
Db 4 GPAGG 8

RESULT 4

Q83083

ID Q83083 PRELIMINARY; PRT; 11 AA.  
AC Q83083;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE P13 mini peptide.  
OS Leucania separata nuclear polyhedrosis virus (LsNPV).  
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
OC Nucleopolyhedrovirus.  
OX NCBI\_TaxID=41714;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96140622; PubMed=8572949;  
RA Wang J.W., Qi Y.P., Huang Y.X., Li S.D.;  
RT "Nucleotide sequence of a 1446 base pair SalI fragment and structure  
RT of a novel early gene of Leucania seperata nuclear polyhedrosis  
RT virus.";  
RL Arch. Virol. 140:2283-2291(1995).  
DR EMBL; U30303; AAA99737.1; -.  
SQ SEQUENCE 11 AA; 1339 MW; F7BDBE0BD40DC401 CRC64;

Query Match 19.0%; Score 4; DB 12; Length 11;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 NRHR 20  
    ||||  
Db 2 NRHR 5

RESULT 5

O03816

ID O03816 PRELIMINARY; PRT; 12 AA.  
AC O03816; O03592;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Chloroplast subunit of light-independent protochlorophyllide reductase  
DE (Fragment).  
GN CHLB GENE.  
OS Metasequoia glyptostroboides.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Metasequoia.  
OX NCBI\_TaxID=3371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97263785; PubMed=9108142;  
RA Karpinska B., Karpinski S., Hallgren J.E.;  
RT "The chlB gene encoding a subunit of light-independent

RT protochlorophyllide reductase is edited in chloroplast of conifers.";  
 RL Curr. Genet. 31:343-347(1997).  
 DR EMBL; X98574; CAA67183.1; -.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1475 MW; 0AB0FB7CC2276724 CRC64;

Query Match 19.0%; Score 4; DB 8; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AWFN 17  
 ||||  
 Db 9 AWFN 12

RESULT 6  
 O03815

ID O03815 PRELIMINARY; PRT; 12 AA.  
 AC O03815; O03591;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Chloroplast subunit of light-independent protochlorophyllide reductase  
 DE (Fragment).  
 GN CHLB GENE.  
 OS Abies alba (Edeltanne) (European silver fir).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.  
 OX NCBI\_TaxID=45372;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97263785; PubMed=9108142;  
 RA Karpinska B., Karpinski S., Hallgren J.E.;  
 RT "The chlB gene encoding a subunit of light-independent  
 protochlorophyllide reductase is edited in chloroplast of conifers.";  
 RL Curr. Genet. 31:343-347(1997).  
 DR EMBL; X98570; CAA67179.1; -.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1475 MW; 0AB0FB7CC2276724 CRC64;

Query Match 19.0%; Score 4; DB 8; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AWFN 17  
 ||||  
 Db 9 AWFN 12

RESULT 7  
 P92454

ID P92454 PRELIMINARY; PRT; 12 AA.

AC P92454;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Chloroplast subunit of light-independent protochlorophyllide reductase  
 DE (Fragment).  
 GN CHLB GENE.  
 OS *Cycas revoluta* (Cycad).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Cycadophyta; Cycadales; Cycadaceae; *Cycas*.  
 OX NCBI\_TaxID=3396;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97263785; PubMed=9108142;  
 RA Karpinska B., Karpinski S., Hallgren J.E.;  
 RT "The chlB gene encoding a subunit of light-independent  
 RT protochlorophyllide reductase is edited in chloroplast of conifers.";  
 RL Curr. Genet. 31:343-347(1997).  
 DR EMBL; X98572; CAA67181.1; -.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 12 12  
 SQ SEQUENCE 12 AA; 1487 MW; 083AFB7163176724 CRC64;  
  
 Query Match 19.0%; Score 4; DB 8; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 Qy 14 AWFN 17  
 ||||  
 Db 9 AWFN 12

RESULT 8  
 Q36669  
 ID Q36669 PRELIMINARY; PRT; 12 AA.  
 AC Q36669;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Chloroplast subunit of light independent protochlorophyllide reductase  
 DE (Fragment).  
 GN CHLB.  
 OS *Pinus sylvestris* (Scots pine).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; *Pinus*.  
 OX NCBI\_TaxID=3349;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Cotyledon;  
 RX MEDLINE=97263785; PubMed=9108142;  
 RA Karpinska B., Karpinski S., Hallgren J.E.;  
 RT "The chlB gene encoding a subunit of light-independent  
 RT protochlorophyllide reductase is edited in chloroplast of conifers.";  
 RL Curr. Genet. 31:343-347(1997).

DR EMBL; X98684; CAA67241.1; -.  
KW Chloroplast.  
FT NON\_TER 1 1  
FT NON\_TER 12 12  
SQ SEQUENCE 12 AA; 1475 MW; 0AB0FB7CC2276724 CRC64;

Query Match 19.0%; Score 4; DB 8; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AWFN 17  
|||  
Db 9 AWFN 12

#### RESULT 9

Q37790

ID Q37790 PRELIMINARY; PRT; 12 AA.  
AC Q37790;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last annotation update)  
DE Chloroplast subunit of light independent protochlorophyllide reductase  
DE (Fragment).  
GN CHLB.  
OS Larix eurolepis.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Larix.  
OX NCBI\_TaxID=49226;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Cotyledon;  
RA Karpinska B., Karpinski S., Hilgren J.E.;  
RL Curr. Genet. 0:0-0(0).  
DR EMBL; X98687; CAA67245.1; -.  
DR EMBL; X98680; CAA67238.1; -.  
KW Chloroplast.  
FT NON\_TER 1 1  
FT NON\_TER 12 12  
SQ SEQUENCE 12 AA; 1475 MW; 0AB0FB7CC2276724 CRC64;

Query Match 19.0%; Score 4; DB 8; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AWFN 17  
|||  
Db 9 AWFN 12

#### RESULT 10

P92460

ID P92460 PRELIMINARY; PRT; 13 AA.  
AC P92460;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Chloroplast subunit of light-independent protochlorophyllide reductase  
 DE (Fragment).  
 GN CHLB GENE.  
 OS *Taxus baccata* (English yew).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Taxaceae; *Taxus*.  
 OX NCBI\_TaxID=25629;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97263785; PubMed=9108142;  
 RA Karpinska B., Karpinski S., Hallgren J.E.;  
 RT "The chlB gene encoding a subunit of light-independent  
 RT protochlorophyllide reductase is edited in chloroplast of conifers."  
 RL Curr. Genet. 31:343-347(1997).  
 DR EMBL; X98571; CAA67180.1; -.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 13 13  
 SQ SEQUENCE 13 AA; 1588 MW; 578AB0FB7CC22767 CRC64;

Query Match 19.0%; Score 4; DB 8; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 AWFN 17  
 ||||  
 Db 9 AWFN 12

# RESULT 11

Q36622

ID Q36622 PRELIMINARY; PRT; 13 AA.  
 AC Q36622;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Chloroplast subunit of light independent protochlorophyllide reductase  
 DE (Fragment).  
 GN CHLB.  
 OS *Picea abies* (Norway spruce) (*Picea excelsa*).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; *Picea*.  
 OX NCBI\_TaxID=3329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Cotyledon;  
 RX MEDLINE=97263785; PubMed=9108142;  
 RA Karpinska B., Karpinski S., Hallgren J.E.;  
 RT "The chlB gene encoding a subunit of light-independent  
 RT protochlorophyllide reductase is edited in chloroplast of conifers."  
 RL Curr. Genet. 31:343-347(1997).  
 DR EMBL; X98685; CAA67242.3; -.  
 KW Chloroplast.  
 FT NON\_TER 1 1

FT NON\_TER 13 13  
SQ SEQUENCE 13 AA; 1585 MW; 4A4AB0FB7CC22767 CRC64;

Query Match 19.0%; Score 4; DB 8; Length 13;  
Best Local Similarity 100.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 AWFN 17  
||||  
Db 9 AWFN 12

#### RESULT 12

Q9UC90

ID Q9UC90 PRELIMINARY; PRT; 15 AA.  
AC Q9UC90;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE ONCOFETAL-laminin binding COLLAGENALPHA 1(I) chain, of-LB collagen  
DE alpha 1(I) (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95169134; PubMed=7864881;  
RA Minafra I.P., Andriolo M., Basirico L., Aquino A., Minafra S.,  
RA Boutillon M.M., van der Rest M.;  
RT "Onco-fetal/laminin-binding collagen from colon carcinoma: detection  
RT of new sequences.";  
RL Biochem. Biophys. Res. Commun. 207:852-859(1995).  
FT NON\_TER 1 1  
FT NON\_TER 15 15  
SQ SEQUENCE 15 AA; 1363 MW; 338D20B78D3B1BF2 CRC64;

Query Match 19.0%; Score 4; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.4e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GPAG 11  
||||  
Db 8 GPAG 11

#### RESULT 13

Q9UC88

ID Q9UC88 PRELIMINARY; PRT; 16 AA.  
AC Q9UC88;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE ONCOFETAL-laminin binding COLLAGENALPHA 1(I) chain, of-LB collagen  
DE alpha 1(I) (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=95169134; PubMed=7864881;  
 RA Minafra I.P., Andriolo M., Basirico L., Aquino A., Minafra S.,  
 RA Boutillon M.M., van der Rest M.;  
 RT "Onco-fetal/laminin-binding collagen from colon carcinoma: detection  
 RT of new sequences.";  
 RL Biochem. Biophys. Res. Commun. 207:852-859(1995).  
 FT NON\_TER 1 1  
 FT NON\_TER 16 16  
 SQ SEQUENCE 16 AA; 1385 MW; F7F93D20B39D3B1B CRC64;

Query Match 19.0%; Score 4; DB 4; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GPAG 11  
 ||||  
 Db 8 GPAG 11

#### RESULT 14

Q9UC54

ID Q9UC54 PRELIMINARY; PRT; 16 AA.  
 AC Q9UC54;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE 18 kDa basic fibroblast growth factor homolog (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=96147281; PubMed=8564983;  
 RA Izbiccka E., Dunstan C., Esparza J., Jacobs C., Sabatini M.,  
 RA Mundy G.R.;  
 RT "Human amniotic tumor that induces new bone formation in vivo produces  
 RT growth-regulatory activity in vitro for osteoblasts identified as an  
 RT extended form of basic fibroblast growth factor.";  
 RL Cancer Res. 56:633-636(1996).  
 SQ SEQUENCE 16 AA; 1431 MW; 5E3B05FFA38A0A09 CRC64;

Query Match 19.0%; Score 4; DB 4; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GPAG 11  
 ||||  
 Db 5 GPAG 8

#### RESULT 15

Q9TWK1

ID Q9TWK1 PRELIMINARY; PRT; 16 AA.  
AC Q9TWK1;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE Proximal collagen (Fragment).  
OS Mytilus edulis (Blue mussel).  
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Mytiloida;  
OC Mytiloidea; Mytilidae; Mytilus.  
OX NCBI\_TaxID=6550;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95230211; PubMed=7714453;  
RA Qin X., Waite J.H.;  
RT "Exotic collagen gradients in the byssus of the mussel Mytilus  
RT edulis.";  
RL J. Exp. Biol. 198:633-644(1995).  
SQ SEQUENCE 16 AA; 1418 MW; 72887820BC282660 CRC64;

Query Match 19.0%; Score 4; DB 5; Length 16;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 GPAG 11  
|||  
Db 2 GPAG 5

Search completed: November 13, 2003, 10:38:20  
Job time : 53.625 secs